

Complex Signals and Perceiver Behavior

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Tricia L. Rubi

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

David W. Stephens

December, 2015

© Tricia L. Rubi 2015
ALL RIGHTS RESERVED

Acknowledgements

First, I would like to thank my colleagues in the Stephens Lab, who have played an invaluable role in helping me develop and complete the work presented in this thesis. Thanks to my advisor, Dave Stephens, for developing these ideas with me, and for teaching me how to be a productive and savvy academic. Perhaps just as importantly, thanks to Dave for always making us laugh, be it at our expense, at his expense, or at the expense of biogeochemists. Tim Polnaszek and Virginia Heinen have been amazing lab mates who have helped me through every stage of this work, from experimental design to manuscript preparation. Thanks to Jonathan Keller and Evan Carter for collaborating with me on this and other work, and for general feedback and assistance. Finally, a special thanks to all of the undergraduate workers and volunteers in the Stephens Lab, without whom this work would be literally impossible.

Many others have contributed in large and small ways to this research. Thanks to my committee members, Mark Bee, Emilie Snell-Rood, Marlene Zuk, and Mike Wilson, for all their advice and feedback on this work. In particular, Mark and Emilie have been instrumental in helping me develop these ideas, and have been a source of both academic inspiration and moral support. Thanks to George Uetz and Dave Clark for collaborating with me on the *Schizocosa* research. Thanks to Ofer Tchernichovski, Esteban Fernández-Juricic, and Will Pearse for generously sharing their expertise.

Thanks to all my funding sources, including the Elmer C. Birney Fellowship, the Carol H. and Wayne A. Pletcher Fellowship, the Alexander and Lydia Anderson Fellowship, the Richard and Judi Huempfer Ruffed Grouse Fellowship, the Animal Behavior Society Student Research Grant, the UMN Council of Graduate Students Travel Fund, the International Society for Behavioral Ecology Travel Fund, and the UMN EEB department for multiple summer research grants and travel grants.

The EEB department at the University of Minnesota was a wonderful place to be a graduate student. Thanks to the entire EEB community, and in particular Sarah Hobbie and Scott Lanyon, for striving to make EEB a welcoming and supportive environment. Thanks to all the behavior folks, in particular the Behavior Group Seminar contingent, for sharing their love of animal behavior. Thanks to all my friends at the U for making these years a fun and inspiring experience.

Thanks to my partner, David Goluskin, for many years of love, support, and grammar criticism. Thanks to my parents for always being my biggest fans and my best friends.

Finally, I would like to thank the blue jays that served as my experimental subjects. Though they weren't given much choice on the matter, they gave up their freedom and their fitness in the name of science, and I genuinely appreciate their sacrifice.

Contents

Acknowledgements	i
List of Tables	vi
List of Figures	vii
1 Introduction	1
1.1 A Note on this Experimental Approach	2
1.2 Complex Signals and Receivers	3
1.3 Complex Signals and Eavesdroppers	6
1.4 Future Directions	7
2 Should receivers follow multiple signal components? An economic perspective.	10
2.1 Introduction	10
2.2 Methods	13
2.2.1 Subjects	14
2.2.2 Testing Apparatus	14
2.2.3 Stimuli	14
2.2.4 Free Trial Walkthrough	17
2.2.5 Trial Overview	17
2.2.6 Treatment Progression and Termination	18
2.2.7 Data Analysis	19
2.3 Results	20
2.4 Discussion	24

2.4.1	Summary of Results	24
2.4.2	Connections to Saliency	25
2.4.3	Alternative Explanations for the Prevalence of Complex Signals	26
2.4.4	Concluding Summary	28
3	Does multimodality <i>per se</i> improve receiver performance? An explicit comparison of multimodal versus unimodal complex signals	29
3.1	Introduction	29
3.2	Methods	32
3.2.1	Subjects	32
3.2.2	Testing Apparatus	32
3.2.3	Stimuli and Treatment Design	33
3.2.4	Trial Walkthrough	34
3.2.5	Trial Types	34
3.2.6	Stability Assessment	35
3.2.7	Data Analysis	35
3.3	Results	37
3.3.1	Performance Measures	37
3.3.2	Component Preference	39
3.3.3	Component Integration	40
3.4	Discussion	40
3.4.1	Summary of Results	40
3.4.2	Alternative Benefits of Multimodal Signals	42
3.4.3	Concluding Summary	43
4	An experimental test of the “Confirm the Rare State” hypothesis	44
4.1	Introduction	44
4.2	Methods	48
4.2.1	Subjects	48
4.2.2	Testing Apparatus	48
4.2.3	Signals	49
4.2.4	Trial Walkthrough	50
4.2.5	Treatment Overview	51

4.2.6	Data Analysis	52
4.3	Results	52
4.3.1	Stable signal following behavior	52
4.3.2	Transition from Pre-treatment ($p = 0.5$)	54
4.4	Discussion	54
4.4.1	Overview of Results	54
4.4.2	Rare States and Signaling Systems	56
4.4.3	Concluding Summary	57
5	Detection of signaling <i>Schizocosa ocreata</i> by an avian predator	58
5.1	Introduction	58
5.2	Methods	61
5.2.1	Subjects	61
5.2.2	Testing Apparatus	61
5.2.3	Video Stimuli	61
5.2.4	Trial Walkthrough	63
5.2.5	Experiment Progression	65
5.2.6	Data Analysis	65
5.3	Results	66
5.3.1	Detectability	66
5.3.2	Reaction Time	67
5.4	Discussion	69
5.4.1	Results in Context	69
5.4.2	Multi-Species Predation and Complex Signals	71
5.4.3	Concluding Summary	72
	References	73

List of Tables

2.1	The values a , b , c , and d are acceptance rates for each of the four stimulus types, representing every possible combination of components. These rates were used to calculate the four scores used as dependent measures in the analysis.	20
3.1	Accept rates to all four possible combinations of components. The values a , b , c , and d are acceptance rates. These values were used to calculate the interaction score to assess component integration.	37
4.1	Accept rates for each of the four component combinations for all five signal following strategies. The final two strategies (Accept Only S+T+, Reject Only S-T-) are the “Confirm the Rare State” strategies. For the Follow One Component strategy, the example of following S only is shown.	46
4.2	Observed accept rates for the four component combinations.	52

List of Figures

2.1	Predicted results. Subjects will follow one component in every treatment. When reliabilities are unequal the more reliable component will be followed; when they are equal, some subjects will follow color, some will follow pattern.	13
2.2	Diagram of the testing apparatus. (a) Overhead view. (b) Front panel. .	15
2.3	Stimulus sets. Each row is a stimulus set. In order from left to right, the stimulus categories are <i>dark/thick</i> , <i>light/thick</i> , <i>dark/thin</i> , and <i>light/thin</i> .	16
2.4	Example learning curve. Daily free trial accept rates for a single bird in a single treatment to each of the four stimulus types. Days 1-3 were the learning phase, days 4-5 were unstable, and days 6-10 showed stable responding. All data used in the analysis were collected from stable responses. This bird is showing pattern-following behavior (accepting any P+, rejecting any P-).	19
2.5	Effect of color and pattern across all treatments. Subjects strongly followed the more reliable component when reliability differences were large. Color following was seen more often than expected; all subjects followed color when it was equally reliable relative to pattern, and individual preferences were seen when color was slightly more reliable (i.e., some subjects followed color and some followed pattern). We attribute this trend to a slight color bias in our subjects (this is explored further in the Discussion section). Note that intermediate effect size values result from pooling data from color followers and pattern followers; they denote a mix of strategies across subjects rather than intermediate follow rates.	21

2.6	Interaction effect across all treatments. The interaction effect did not differ significantly between treatments.	22
2.7	Smaller effect versus larger effect. The gray bars show the size of the effect of the “less followed” component (regardless of whether that component was color or pattern). This effect was approximately zero in every treatment and did not differ significantly between treatments. The white bars show the size of the larger effect for comparison.	23
3.1	Example stimuli. We refer to component 1 as the fixed component, since it is a color in both unimodal and multimodal treatments. We refer to component 2 as the varied component, since it is a color in unimodal treatments and a tone in multimodal treatments. (a) A unimodal stimulus: component 1 is the forewing color and component 2 is the hindwing color. (b) A multimodal stimulus: component 1 is the color of the wings and component 2 is a pure tone.	31
3.2	An example configuration of the testing apparatus (the left-right positions of the accept and reject keys were randomized across subjects) (a) Overhead view. (b) Front panel view.	33
3.3	Proportion of correct responses across all treatments. A repeated measures ANOVA revealed a significant interaction between the reliability of the fixed component and the reliability of the varied component. There were no significant effects of modality or interactions involving modality.	38
3.4	Number of trials until stability criterion was reached. A repeated measures ANOVA revealed a significant main effect of the reliability of the varied component. There were no significant effects of modality or interactions involving modality.	38
3.5	Proportion of “Follow Varied Component” responses when the two stimuli disagreed. A repeated measures ANOVA revealed significant main effects of modality, the reliability of the fixed component, and the reliability of the varied component.	39
3.6	Interaction scores across all treatments. The black dots indicate means, lines indicate medians, box edges span the first to third quartiles, whiskers extend to 1.5 * interquartile range, + indicates an outlier.	40

4.1	Payoffs for the five strategies across a range of p	47
4.2	An example configuration of the testing apparatus (left-right assignments for the accept and reject keys were randomized across subjects). (a) Overhead view. (b) Front panel.	49
4.3	The two signal sets. The positive / negative assignments were randomized such that each combination was the S+T+ for at least one subject. . . .	50
4.4	Box plot of the strategy Scores in the two treatments. The black dots indicate means, lines indicate medians, box edges span the first to third quartiles, whiskers extend to 1.5 * interquartile range, triangles indicate outliers. The expected values in each treatment are indicated with \otimes . . .	53
4.5	Accept rates for the four component combinations. Data are grouped into blocks of 150. These plots show the last 3 blocks of the pre-treatment ($p = 0.5$) and the first nine blocks of the treatments. The dashed lines separate the pre-treatments from the treatments. (a) Good state – rare treatment ($p = 0.2$). (b) Bad state – rare treatment ($p = 0.8$).	55
5.1	An example schematic of the testing apparatus. The positions of the “present” and “absent” keys were randomized across subjects. (a) Overhead view. (b) Front panel.	62
5.2	Example stationary spider stimuli. Courting and walking spiders looked very similar, but the spider was performing the appropriate motion. Neutral gray backgrounds were used in the pre-experiment training (panels a and b), and leaf litter backgrounds were used in the experiment (panels c and d). Blank spider-absent stimuli looked identical, but with no spider present. The spider is located to the left of the arrow. (a) A natural-colored spider on a neutral gray background. (b) A background-matching spider on a neutral gray background. (c) A natural-colored spider on a leaf litter background. (d) A background-matching spider on a leaf litter background.	64
5.3	Detectability (measured by d') in the last three blocks. Bk-Matching stands for background-matching.	67

5.4	Detectability (measured by d') of all stimulus types over the course of the entire experiment. Bk-Matching stands for background-matching coloration.	68
5.5	Reaction time for courting and walking spiders in the last three blocks when the response was a hit. Bk-Matching stands for background-matching.	68

Chapter 1

Introduction

A male Temminck's tragopan woos a prospective mate by performing an elaborate dance while displaying a brilliant, multicolored wattle and bright blue, fleshy horns; a garden tiger moth warns a prospective predator of its toxicity by flashing bright orange, blue-spotted underwings and emitting an audible stridulation; an eland bull broadcasts his fighting ability via a prominent dewlap, a tuft of dark facial hair, and a loud clicking produced by a tendon in his knee joints. Such elaborate, multi-component displays, called complex signals, are prevalent in animal communication. Complex signal components can range from tightly integrated qualities such as the structural and pigmentary properties of a single color patch (Grether et al. 2004) to elaborate displays spanning several sensory modalities. As our ability to observe and characterize signal attributes improves, an increasing number of researchers are arguing that complex signals are the rule, rather than the exception, in animal communication (Hebets and Papaj 2005; Partan and Marler 1999; Rowe et al. 2004; Wilson et al. 2013). They are observed across a wide range of taxa and have been reported in every major type of signaling interaction, including courtship (e.g., Costanzo and Monteiro 2007; Elias et al. 2005; Grafe et al. 2012; Smith and Evans 2008; reviewed in Candolin 2003), aposematism (e.g., Claridge 1974; Hauglund et al. 2006; Prudic et al. 2006; reviewed in Rowe and Guilford 1999), agonism (e.g., Bro-Jørgensen and Dabelsteen 2008; Deag and Scott 1999; Hughes 1996), begging (e.g., Kim et al. 2011; Leonard et al. 2003; Stynoski and Noble 2011), and individual identity signals (e.g., DelBarco-Trillo et al. 2012; Kondo et al. 2012; Proops et al. 2008; Tibbetts 2002). The striking prevalence of complex signals in nature suggests that

they are broadly beneficial for signalers, however, we still have much to learn about how signal components interact and the implications for communication.

The broad aim of this thesis is to explore complex signaling from the perspective of perceivers – the animals that detect and respond to signals. Complex signals are likely to pose particular challenges and benefits to perceivers simply as a result of having multiple components. For example, classic learning studies have identified several ways in which multiple stimuli can interfere with each other (Mackintosh 1975; Pavlov 1927; Rescorla and Wagner 1972), as well as ways in which multiple stimuli can enhance each other (Clarke et al. 1979; Couvillon and Bitterman 1982). Complex signals are also thought to be more detectable to intended and unintended perceivers alike (Candolin 2003; Hebets and Papaj 2005; Rowe 1999). Complex signals may have components in multiple sensory modalities (e.g., visual, acoustic, olfactory), which itself may have special implications for detection and processing. In the following chapters, I investigate how perceivers utilize the various components of complex signals, and consider the implications of perceiver behavior for complex signaling as a signaler strategy. In chapters 2 - 4, I focus on receivers – the intended recipients of communicative signals. In chapter 5, I shift my focus to eavesdroppers – unintended signal recipients that exploit signals, often to the detriment of signalers. In each of these experiments, captive blue jays (*Cyanocitta cristata*) serve as my experimental perceivers.

1.1 A Note on this Experimental Approach

These experiments are lab-based, operant learning tasks in which I present subjects with artificial or digitized signals and use button-pecking to characterize signal following behavior; clearly, this experimental system is quite divorced from natural signaling systems. Nevertheless, this approach is a powerful tool with which to investigate the costs and benefits that shape animal behavior. Testing learned responses to artificial signals removes (or at least minimizes) the influence of prior biases, providing a clearer picture of the economic forces shaping behavior. In these experiments, I was able to manipulate and quantify signal attributes and comprehensively describe the signal following strategies adopted by receivers. This level of control allowed me to test complex signaling hypotheses in a manner that would be difficult or impossible to achieve in

natural communication systems. While these experiments do not explicitly test evolutionary hypotheses, conducting quantitative analyses of costs and benefits in learned signal following tasks may provide key insights on the evolutionary pressures shaping communication in nature. As such, I believe that this research fills a vacant niche in the communication literature. My hope is that this work will help to develop a framework that can be used to guide future research in complex signaling, including studies in more naturalistic communication systems.

1.2 Complex Signals and Receivers

The function of a signal is to influence receivers to behave in a way that is beneficial to signalers; our courting tragopan benefits when a female mates with him, and our garden tiger moth benefits when a predator avoids it. A signal that is ignored or undetected by receivers is functionally useless, therefore, it is critical to consider the selective pressures arising from receiver responses. As mentioned previously, the near ubiquity of complex signals suggests that they are beneficial for signalers, and a popular hypothesis is that complex signals are beneficial because they are more effective at influencing receiver behavior. Researchers have identified many perceptual and psychological benefits of complex signals, such as improved detection or enhanced memorability (e.g., Candolin 2003; Hebets and Papaj 2005; Rowe 1999). More traditional economic approaches have also identified several scenarios that could favor complex signaling, including specialized cost structures (Iwasa and Pomiankowski 1994; Moller and Pomiankowski 1993; Pomiankowski and Iwasa 1993; Wilson et al. 2013), multiple receiver preferences (Andersson et al. 2002; Ratcliffe and Nydam 2008; Wilson et al. 2013), or “multiple message” signals (signals in which components convey subtly different information about an overarching quality – Johnstone 1996; Partan and Marler 1999; Wilson et al. 2013). Despite these successes, several researchers have commented on a persistent, somewhat unexpected result: from a purely informational perspective, it is unclear why signals with multiple components should be any better at conveying a message than a single signal is (e.g., Bro-Jørgensen 2010; Johnstone 1995; Schluter and Price 1993; Wilson et al. 2013). In the absence of additional constraints, multiple signal components offer no obvious informational benefits over a single component. Though this observation does

not preclude complex signaling as a valid strategy, it is surprising given the abundance of complex signals in communication. To understand the broad benefits of complex signals, we must first understand when complex signal use is predicted to be economically advantageous and when it is not. Chapters 2 - 4 focus on the economics of complex signaling from the receiver perspective. For simplicity, I focus on cases in which the interests of the signaler and receiver are aligned; that is, both parties benefit when the receiver's response matches the signaler's true underlying state (e.g., when a predator avoids a toxic prey item). In such systems, receiver performance can be used to characterize the effectiveness of the signal.

Chapter 2 tackles a fundamental, yet previously untested question: do receivers benefit economically from following a reliable, redundant complex signal? I develop and test a simple economic model that predicts (like others before it) that receivers will not utilize multiple signal components in a basic signal. I presented receivers with signals with two components, a color and a pattern, each of which was independently informative about an underlying state. I also experimentally varied component reliability, defined as the degree to which a component accurately indicated the underlying state. Consistent with my predictions, I found that receivers never integrated two components into decision-making. Instead, they selected a single component and followed only that one. This study confirms that we should not assume that receivers will follow all seemingly important signal attributes. In addition, the study outlines a scenario in which following multiple components is not a beneficial strategy, an important first step in understanding when such behavior *is* beneficial.

Chapter 3 tests the hypothesis that multimodal signals, defined as complex signals with components in several sensory modalities, are more effective at influencing receiver behavior than unimodal signals are (Partan 2013; Rowe and Halpin 2013; Siddall and Marples 2008). This is a popular hypothesis with broad experimental support, yet almost all previous studies have compared whole multimodal signals to their isolated components (e.g., Hughes 1996; Narins and Grabul 2005; Siddall and Marples 2008; Stynoski and Noble 2011; Uetz et al. 2009; VanderSal and Hebets 2007, but see Hauglund et al. 2006; Kulahci et al. 2008; Taylor and Ryan 2013). Comparing whole multimodal signals to constituent components confounds the effects of multiple modalities with the effects of multiple components. This is problematic for two reasons. First, most of these

studies focus on natural signals; responses to partial signals (relative to whole signals) might be reduced even in the absence of explicit modality effects. Second, many of the proposed benefits of multimodal signals have been attributed to complex signals in general (Hebets and Papaj 2005). In this experiment, I compared receiver responses to two-component signals that differed only in whether they were cross-modal. Based on several measures of performance, I found no significant differences in how receivers followed multimodal versus unimodal signals. I also found no evidence that receivers are more likely to integrate multiple components into decision-making when signals are multimodal. In fact, I found that a single signal component dominated decision-making, consistent with the model developed in chapter 2. These findings challenge the generality of the claim that multimodality is inherently beneficial for receivers. At the end of this chapter I briefly explore some conditions under which multimodal signaling might be beneficial.

In chapter 4, I test a new hypothesis developed in collaboration with my advisor, Dr. David Stephens. In a separate work (Rubi and Stephens *in revision*), we developed a model that illustrated that receivers can sometimes benefit economically from following two signal components. We found that when one signaler state is less common (e.g., “high quality male,” “toxic prey”), the optimal receiver strategy is to use both signal components to determine whether the signaler is in the rare state. (As a side note, we also found that following two components is never better – and is often worse – than following one component when signaler states are balanced; therefore, these findings are consistent with the results presented in chapters 2 and 3, which tested cases in which the good and bad states were equally common.) The experiment presented in chapter 4 is the first experimental test of the model developed in Rubi and Stephens (*in revision*). I focused on two experimental cases, one in which the “good” signaler state was rare, and one in which the “bad” signaler state was rare, and characterized the signal following strategies the receivers adopted. The results of this experiment were somewhat mixed; I found that receiver behavior followed the basic pattern that we predicted, but the trend was weak. While this result suggests that our model has merit in predicting complex signal use by receivers, more work needs to be done in this area.

1.3 Complex Signals and Eavesdroppers

Eavesdroppers can exploit signals in a variety of ways, many of which are harmful to signalers. For example, males may use the courtship calls of rivals to intercept females (e.g., Milner et al. 2010), or heterospecific competitors may eavesdrop on food advertisement calls directed at conspecific groupmates (e.g., Nieh et al. 2004). In this thesis I focus on eavesdropping predators, which can impose particularly dire consequences on signalers. Signals are selected to be highly detectable (Guilford and Dawkins 1991, 1993; Wiley 2006), and as a result, signaling animals are often especially vulnerable to predators. It is well established that predators exploit the signals of their prey, and that signalers face a trade-off between the benefits of successful communication and the costs of increased conspicuousness to predators (Magnhagen 1991; Zuk and Kolluru 1998). This trade-off is likely to be even more pronounced for complex signals; enhanced detectability is one of the most commonly-cited benefits of signals with multiple components (Candolin 2003; Hebets and Papaj 2005; Partan and Marler 2005; Rowe 1999). Multimodal signals in particular may incur high predation costs simply because they will be perceptible by a wider range of predator species. Despite the logic of these arguments, the literature is currently lacking an explicit discussion of complex signals and multi-species predation.

The experiment presented in chapter 5 was completed in collaboration with Drs. George Uetz (University of Cincinnati) and Dave Clark (Alma College) and is part of a broad and comprehensive research program on the costs and benefits of complex sexual signaling in a focal organism, the brush-legged wolf spider (*Schizocosa ocreata*). The courtship display of *S. ocreata* is an elaborate multimodal signal that incorporates several visual and seismic components. Dr. Uetz and his research team have described eavesdropping on the sexual display of *S. ocreata* by several predator species from diverse taxonomic groups. This experiment was the first study examining detection by avian predators, which differ markedly in sensory abilities and predation behavior from the previous predators tested. In this study, I used video playbacks to measure the detectability of signaling and non-signaling spiders. I found that the jays were highly sensitive to motion; walking or courting spiders were much more detectable than stationary spiders. Courting spiders were detected at similar rates to walking spiders, but

were detected more quickly. Taken together with previous work, this study suggests that different predator types exploit different signal components, supporting the hypothesis that complex signalers are especially vulnerable to eavesdropping predators.

1.4 Future Directions

This work suggests a number of potential future directions. In all future studies, it is important to confirm that ostensibly important signal components are indeed influencing receiver behavior in some way (as is illustrated by the results presented in chapters 2 and 3). Several authors have warned researchers against overlooking signal components (e.g., Hebets and Papaj 2005; Partan and Marler 2005), however, it is also possible to make the opposite error. Many studies have found that receivers ignored seemingly significant signal characteristics (e.g., Amorim et al. 2013; Aronsson and Gamberale-Stille 2008; Hebets et al. 2013; Marples et al. 1994; Page and Jaeger 2004; Yorzinski et al. 2013). As these studies show, we cannot assume that traits that are conspicuous to researchers will be important components of a signaling display.

A clear starting point is to continue methodically testing the hypothesized benefits of complex signals for receivers. A particularly compelling hypothesis is that complex signals improve detection in noisy environments (Candolin 2003; Hebets and Papaj 2005; Partan and Marler 2005; Rowe 1999). One possible future study could examine receiver performance when following multimodal signals in various types of environmental “noise,” such as low light conditions and acoustic background noise. If receivers do indeed follow multiple components in such studies, it would be interesting to see how flexible that behavior is. Receivers might show facultative following of the less-inhibited component. Alternatively, receivers trained to follow signals under noisy conditions might continue to follow multiple components even in the absence of noise (suggesting that learning signals in noise is a distinct process). Another popular hypothesis is that signal components might convey subtly different information about a broad quality (the “multiple messages” hypothesis - Johnstone 1995; Moller and Pomiankowski 1993). This hypothesis in particular might benefit from testing with artificial signals because the specific meaning of components is determined by the experimenter.

As discussed in chapter 4, I believe it is worth exploring the relationship between

complex signaling and unbalanced signaler states further. The model developed in Rubi and Stephens (*in revision*) predicts that receivers can benefit economically from following multiple signal components when they are used to confirm that a signaler is in a rare state. This suggests that complex signaling is a beneficial strategy in signaling systems in which one signaler type is rare. Such systems are likely to be common in nature, thus, this is a compelling avenue for future research. The experiment presented in chapter 4 found only weak evidence supporting the predictions of the model, however, receiver behavior roughly followed the predicted trends. Future work could expand on this specific experimental approach, or could look more broadly across natural communication systems for a relationship between signaler frequencies and signal complexity.

As I mention in chapter 5, there is a notable gap in the literature on complex signals and multi-species predation. Signals with multiple components, and in particular multimodal signals, will be perceptible to a wider range of predator species. When predators are the intended receivers (i.e., in aposematic signaling systems), this effect could be highly beneficial. A few researchers have noted that multi-species predation might favor multimodal aposematic signaling (Ratcliffe and Nydam 2008; Rowe and Halpin 2013), but empirical studies are lacking (though see Olofsson et al. 2012). When predators are not the intended recipients, multimodal signals are likely to be highly costly because they can be exploited by a wider range of predator species. There is an extensive body of work on enhanced detectability in complex and multimodal signals (Candolin 2003; Heberts and Papaj 2005; Partan and Marler 2005; Rowe 1999), as well as on predator exploitation of prey signals (Haynes and Yeorgan 1999; Magnhagen 1991; Zuk and Kolluru 1998); however, the overlap between these topics is almost entirely unexplored.

Finally, this work has revealed a somewhat more vague, but noteworthy trend: our subjects had a tendency to minimize the number of signal components they followed. When developing these experiments (as well as in previous work in our lab), we often observed that our subjects preferred “non-signal following” strategies (e.g., always pecking the same button) over signal following strategies. They often adopted these simpler strategies even when they were slightly disadvantageous. For example, when a single-component signal was only slightly reliable (e.g., 60% reliable), we found that many subjects adopted a non-signal following strategy (for an average payoff of 50%).

Essentially, they paid a small cost (10%) to not follow a signal. Similarly, these experiments tested several cases in which following two signal components was equally as good as following a single signal component, and we found that receivers never followed two components in those cases. Unlike the previous example, this behavior was not inconsistent with economic predictions, however, the lack of variability across subjects suggests a bias for following one component instead of two. In other words, there appears to be a slight lag from our economic predictions before receivers switch from following no components, to one component, to two components.

This behavior seems to suggest that component following is somehow intrinsically costly for receivers. Several researchers have suggested that signal following may be costly (e.g., Bradbury and Vehrencamp 2000; Dawkins and Guilford 1991; Jennions and Petrie 1997). Many costs, such as search costs or increased vulnerability to predators, are not specifically applicable to our lab system; nevertheless, the non-following bias we observe is consistent with the general argument that signal following is costly for receivers. In addition, several theoretical models of complex signaling have considered the effect of individual component assessment costs (e.g., Iwasa and Pomiankowski 1994; Moller and Pomiankowski 1993; Pomiankowski and Iwasa 1993); to my knowledge, there is no empirical work in this area. Though determining the mechanistic basis of this phenomenon might be challenging, an operant approach could be used to describe it quantitatively. Assessing the cost of following a single-component signal would be straightforward. Researchers could simply titrate signal reliability (q) upward incrementally until receivers begin following the signal. In the experimental setup described in chapters 2 and 3, the estimate of the cost of following the signal would simply be $q_{signal} - 0.5$. Assessing the costs of following additional signal components would be more complicated, since researchers would need to create a scenario in which following two components is economically advantageous. An approach similar to the one developed in chapter 4 would be appropriate, though work would be needed to determine the ideal parameters to compare single-component strategies and two-component strategies.

Chapter 2

Should receivers follow multiple signal components? An economic perspective.¹

2.1 Introduction

Complex signals, defined as signals consisting of multiple components, are prevalent in animal communication. Indeed, many researchers have suggested that most signals incorporate multiple components (Harper 2006; Hebets and Papaj 2005; Hebets et al. 2013; Rowe and Skelhorn 2004). The near ubiquity of complex signals likely indicates that they are beneficial for signalers, and a commonly proposed benefit is that complex signals are more effective at influencing receiver behavior. A number of empirical studies have supported this claim by showing that multiple components elicit enhanced or improved receiver responses (Akre and Ryan 2010; Kelly and Marples 2004; Kulahci et al. 2008; Leonard et al. 2011; Rowe 2002; Siddall and Marples 2008; Uetz et al. 2009; VanderSal and Hebets 2007), though it is not uncommon to find that receivers respond equivalently to isolated components (Costanzo and Monteiro 2007; Smith and Evans

¹ This chapter is currently in press as Rubi, T.L. and Stephens, D.W. (2015). Should receivers follow multiple signal components? An economic perspective. *Behavioral Ecology*. doi: 10.1093/beheco/arv121.

2008) or prioritize certain components (Amorim et al. 2013; Aronsson and Gamberale-Stille 2008, 2012; Hebets et al. 2013; Marples et al. 1994; Uetz et al. 2009; Yorzinski et al. 2013).

Most of the studies listed above focus on the non-economic benefits of complex signal following (such as improved learning rates or memorability). This is unsurprising given that economic models of receiver behavior predict that receivers should not follow multiple redundant components under basic conditions (e.g., Bro-Jørgensen and Dabelsteen 2008; Iwasa and Pomiankowski 1994; Johnstone 1995; Pomiankowski and Iwasa 1993; Schluter and Price 1993; Wilson et al. 2013). Instead, economic models of stable signal following behavior typically predict that receivers will follow the single most reliable component. It is widely accepted that signal reliability impacts receiver behavior (e.g., Zahavi 1993), however, no empirical studies (to our knowledge) have directly tested the role of component reliability in complex signal following by receivers. This seems like a critical first step; to truly understand when and how complex signals benefit receivers, we must first determine whether economic benefits alone are sufficient to explain preferences for multiple components. In this study we develop a simple economic model examining receiver behavior when stably following a signal with two reliable components, then test it experimentally with live receivers (captive blue jays, *Cyanocitta cristata*). Our aim is to determine whether receivers will incorporate multiple signal components into decision-making under a basic economic framework.

Our development follows the basic structure of the flag model (see McLinn and Stephens 2006; Stephens and Dunlap 2009). Imagine that a signaler produces a signal about some true state, which may be “good” or “bad” (e.g., high quality mate or low quality mate, palatable prey or poisonous prey, etc.). The receiver must choose between two actions, which we call “accept” and “reject.” The receiver should accept when the true state is good and reject when the true state is bad. We assume that the true state is good half the time, and bad half the time, so that a receiver who acts without a signal (e.g., one who guesses) will choose correctly 50% of the time. First, consider a single-component signal (S), for example, a color. The signal can take two forms, S+ (e.g., red) or S- (e.g., green). The signal is useful to the receiver because the two forms are statistically associated with the true state; S+ is associated with the good state, and S- is associated with the bad state. We measure the reliability of this association using

the conditional probability q , which ranges between 0.5 and 1. If $q = 1.0$ the signal form is perfectly correlated with the true state, but if $q = 0.5$ the signal's information is no better than chance. When $q > 0.5$, the receiver will perform better if it follows the signal (that is, if it accepts when it observes S+ and rejects when it observes S-).

Now we extend this scenario to a two-component signal. Consider a signal that consists of a color (C) and a pattern (P). The components independently indicate the true state with reliabilities q_{color} and $q_{pattern}$, respectively. Both components are always visible in a signal, resulting in four possible signals: C+P+, C+P-, C-P+, C-P-. Subjects follow color by accepting a C+ or rejecting a C-, and follow pattern by accepting a P+ or rejecting a P-. When the components agree it doesn't matter which component the subject follows, but when the components disagree receivers should always follow the more reliable component. A receiver that only follows the more reliable component will always do as well or better than a receiver that incorporates both. When the signals are exactly equally reliable a receiver can follow either component, but there is no advantage to following both. We predict therefore that an optimal signal follower will follow color if the reliability of color exceeds the reliability of pattern, and vice versa. Figure 2.1 shows this prediction graphically. We predict color following in the triangular region above the diagonal, and pattern following below the diagonal. Notice that the model makes no explicit predictions about what should happen along the diagonal ($q_{color} = q_{pattern}$); following color and following pattern are equivalent strategies. In a pilot study, we observed that when reliabilities are equal some subjects prefer to follow color and some to follow pattern, so in these treatments we expect to see a mix of strategies across subjects.

We tested this model in an experiment using captive blue jays (*Cyanocitta cristata*) with artificial signals composed of two components. Our experiment exploited the jays' ability to learn the relevant properties of signals and adjust their responses to signals accordingly. We used butterfly-shaped stimuli for our experimental signals; this choice was motivated by the well-known aposematic signaling system that exists between blue jays and monarch butterflies (Brower 1969). The butterfly stimuli differed only in coloration and patterning of the wings, representing a signal with two visual components (color and pattern). We experimentally manipulated color and pattern reliabilities, so that in some treatments color was more reliable than pattern, and vice versa (Figure

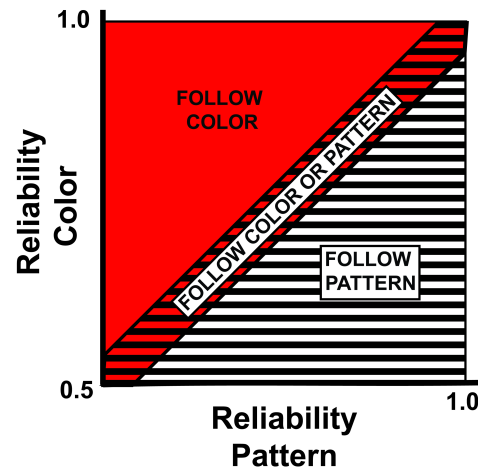


Figure 2.1: Predicted results. Subjects will follow one component in every treatment. When reliabilities are unequal the more reliable component will be followed; when they are equal, some subjects will follow color, some will follow pattern.

2.1). As explained above, we expected subjects to: 1) follow a single component in every treatment, and 2) follow the more reliable component in a given treatment. When color and pattern were equally reliable we expected some individuals to follow color and some to follow pattern.

2.2 Methods

We tested all combinations of three different levels of component reliability: $q = 0.6$, $q = 0.8$, and $q = 1.0$ in a 3×3 factorial design. This resulted in nine total treatments: three treatments in which color was more reliable than pattern ($q_{color} > q_{pattern}$), three treatments in which pattern was more reliable than color ($q_{color} < q_{pattern}$), and three treatments in which color and pattern were equally reliable ($q_{color} = q_{pattern}$). At the end of each treatment, we quantified the degree to which subjects were responding to color, pattern, and combinations of color and pattern. All subjects completed all treatments in a different, randomly chosen order, and comparisons across treatments were made using repeated measures analyses.

2.2.1 Subjects

Eight adult blue jays of unknown sex, randomly selected from a captive colony, served as the test subjects (band numbers 10, 11, 24, 86, 206, 207, 208, and 350). During the experiment subjects were housed individually in the testing chambers for 23 hours / day (they were removed for an hour for cleaning and maintenance). Subjects ran trials from 7 am to 3 pm daily for the duration of the experiment, typically around 12 weeks. They received water *ad libitum* and were maintained on a 12 hour light:dark cycle. The experiment operated in a closed economy; that is, subjects received all of their food from the experiment. To ensure our subjects' well being, we provided additional food in cases where a bird's body weight dropped below 85% of its measured *ad libitum* weight.

2.2.2 Testing Apparatus

Figure 2.2a depicts an overhead view of the testing apparatus. A perch lever was located at the rear of the box directly below a standard pigeon key (MED Associates ENV-123AM). At the front of the box (Figure 2.2b), stimuli were presented on an LCD screen (Accelelevision LCDP7W) located behind a transparent responding apparatus, a pane of Plexiglass wired to a response mechanism. When the apparatus presented an image on the screen, subjects could either accept it by pecking the stimulus itself or reject it by pecking a pigeon key located to the right of the stimulus screen. The food cup, equipped with a magazine light, was located between the accept and reject keys. A single stationary perch spanning the width of the box allowed subjects to access the two keys and the food cup. A Med-PC computer program (Med Associates, Burlington, VT, USA) controlled inputs and outputs and recorded data.

2.2.3 Stimuli

We created nine sets of stimuli, one for each treatment, using Adobe Illustrator (Adobe Systems, San Jose, CA, USA). Each set had a different hue (e.g., red or green) spaced at equal intervals around the Adobe Illustrator color wheel, and a unique line pattern. Within a stimulus set, color had two possible values, light or dark, corresponding to differences in brightness (e.g., bright red or dark red). Pattern had two possible values, thick lines or thin lines. These stimuli were chosen to standardize differences across

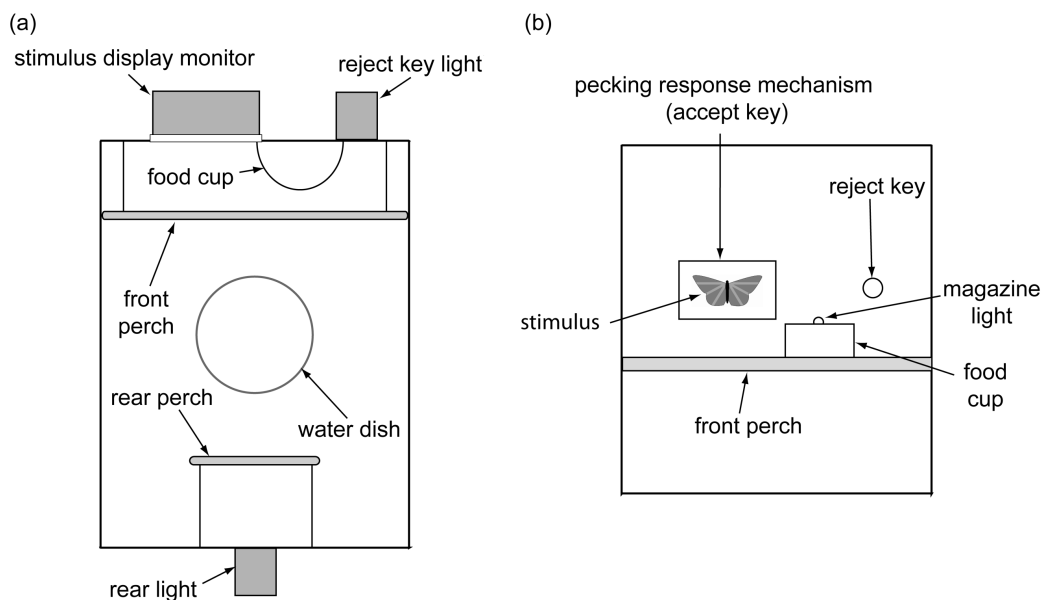


Figure 2.2: Diagram of the testing apparatus. (a) Overhead view. (b) Front panel.

treatments; differences between light/dark and thick/thin were the same in each treatment (brightness was standardized using the Adobe Illustrator Live Color tool). Each stimulus had both a color and a pattern, resulting in four possible stimuli: light/thin, light/thick, dark/thin, and dark/thick. Figure 2.3 shows all nine stimulus sets. (For simplicity, throughout this chapter we refer to brightness and thickness cues as “color” and “pattern,” respectively.)

Each subject completed all nine treatments in a different order and we analyzed the results using repeated measures ANOVA. In a given treatment, each subject experienced a different stimulus set, and subjects experienced each stimulus set once during the course of the experiment. For each subject in each treatment, we randomly assigned light and dark as the positive or negative color state (hereafter C+/C-), and thick and thin the positive or negative pattern state (hereafter P+/P-). We constrained this randomization such that each subject experienced each possible C+P+ combination (e.g., light/thick) at least twice.

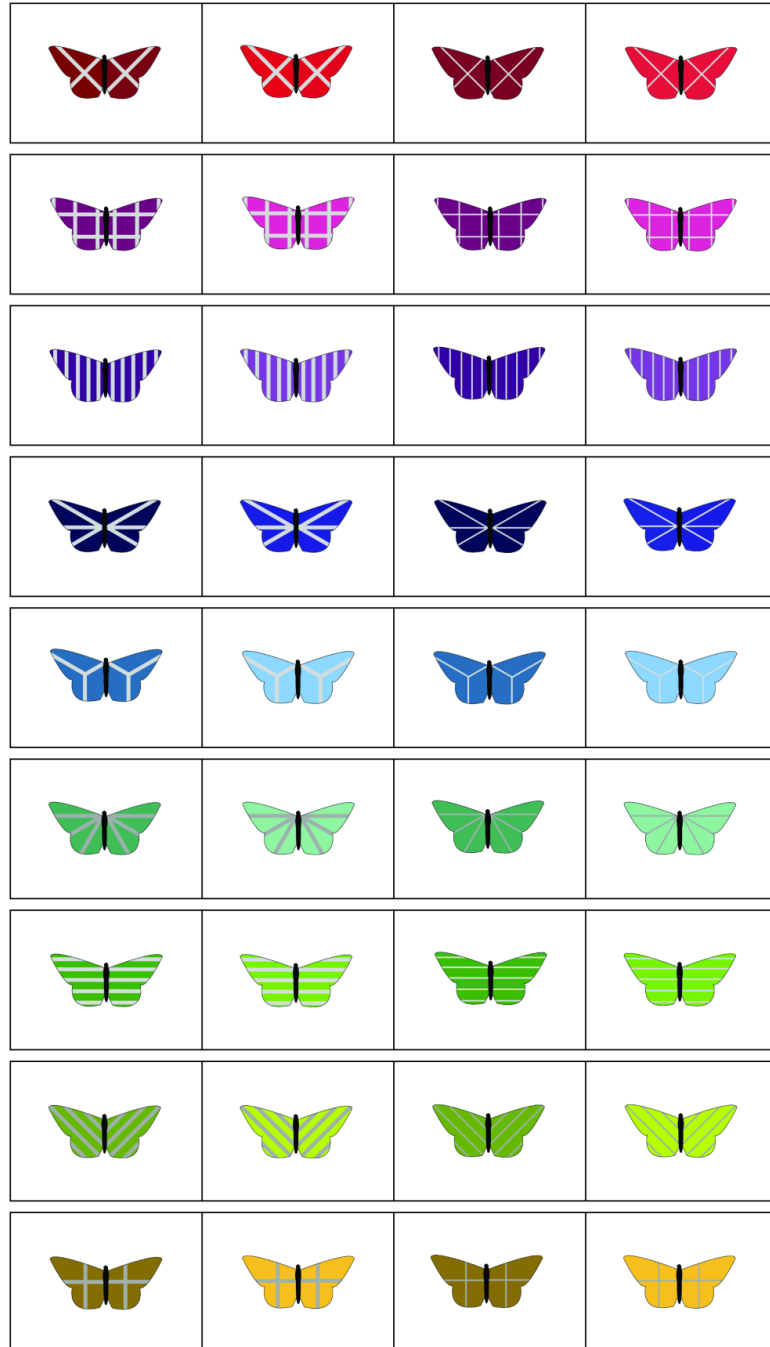


Figure 2.3: Stimulus sets. Each row is a stimulus set. In order from left to right, the stimulus categories are *dark/thick*, *light/thick*, *dark/thin*, and *light/thin*.

2.2.4 Free Trial Walkthrough

Trials were separated by an intertrial interval (ITI) of 90 seconds. At the start of each free trial, the computer determined whether the true state would be “good” or “bad.” The probability that the true state was good was 0.5, so half of the time the rewarded action was accept and half of the time the rewarded action was reject. The computer then independently assigned the forms of the color and pattern components based on the component reliabilities. To begin the trial, a flashing light at the rear of the box indicated that the trial was ready to begin. Once the subject hopped to the rear perch, the flashing light extinguished and the stimulus and reject key illuminated at the front of the box. Subjects responded by either accepting or rejecting the stimulus as described in the Testing Apparatus section. Once the subject responded, the computer extinguished the stimulus and reject key and delivered the programmed food reward, accompanied by a flashing magazine light. The apparatus rewarded correct responses with three food pellets (approximately 0.06 g) and delivered no food for incorrect responses. If a subject did not respond to a stimulus in 15 minutes, the trial aborted, the ITI restarted, and the trial was repeated.

2.2.5 Trial Overview

The first 400 trials of each treatment constituted the learning phase, which allowed time for subjects to learn the treatment parameters. We did not use these trials in the analysis. During the learning phase, subjects experienced two types of trials, forced and free. Forced trials (also known as “no choice” trials) are similar to free trials, except that they require subjects to respond in a predetermined way and experience the associated reward; this ensures that subjects experience all possible choices and outcomes in a treatment. For example, in a forced correct accept trial, the computer assigned the stimulus and the apparatus illuminated the stimulus and reject key, exactly as in free trials. However, the trial only ended when the subject pecked the stimulus (that is, the reject key was not responsive). The subject received a reward in forced correct trials, but not forced incorrect trials. Trials were grouped into blocks of 40, and the first eight trials of every block were forced trials: two correct accepts, two incorrect accepts, two correct rejects, and two incorrect rejects, presented in random order. The remaining 32

trials in a block were free trials, in which subjects responded freely and were rewarded according to responses. The total number of trials completed in the treatment was calculated at the end of each day. When this number exceeded 400, the learning phase ended and the subject advanced to the data collection phase of the treatment.

In order to maintain the proper conditional reliabilities of the components, some compound stimuli (namely, C+P+ and C-P-) occurred more frequently in free trials. The most extreme case was the treatment in which both color and pattern were 100% reliable; in this treatment, free trials could only be C+P+ or C-P- since both components indicated the correct response with perfect reliability. This extreme treatment necessitated the use of probe trials, defined as unrewarded trials presented at a low frequency (in this experiment, 7.5% of all trials). These trials consisted of all four component combinations, delivered at equal frequencies. We introduced probe trials during the data collection phase, and all dependent measures in this experiment were calculated from probe trial responses. The block layout was the same as in the learning phase, except that free trials were randomly replaced with probe trials. Each treatment continued until fourteen probe trials of each component combination were obtained within a period of stable signal following.

2.2.6 Treatment Progression and Termination

After the learning phase had ended, each treatment continued until we collected 14 responses to each type of probe trial (C+P+, C+P-, C-P+, C-P-) during a period of stable responding. Once all 56 probe trials were observed (typically over 4-5 experimental days), we assessed the free trials from that period to determine if behavior was stable. Specifically, we calculated the mean accept rates to each of the four stimulus types over all days, and determined if the daily accept rates differed from the overall accept rates by less than 25%. This indicated that accept rates had reached an asymptote for all four stimulus types. If so, we considered responses stable and ended the treatment. If not, we allowed subjects to continue in the treatment until all probes were collected from a period of stable responding. Figure 2.4 shows an example of this progression (a bird showing pattern-following); behavior changes dramatically during the learning phase (days 1-3), and then begins stabilizing during days 4-5. Days 6-10 fulfilled the stability criterion, and responses from those days were used in the analysis.

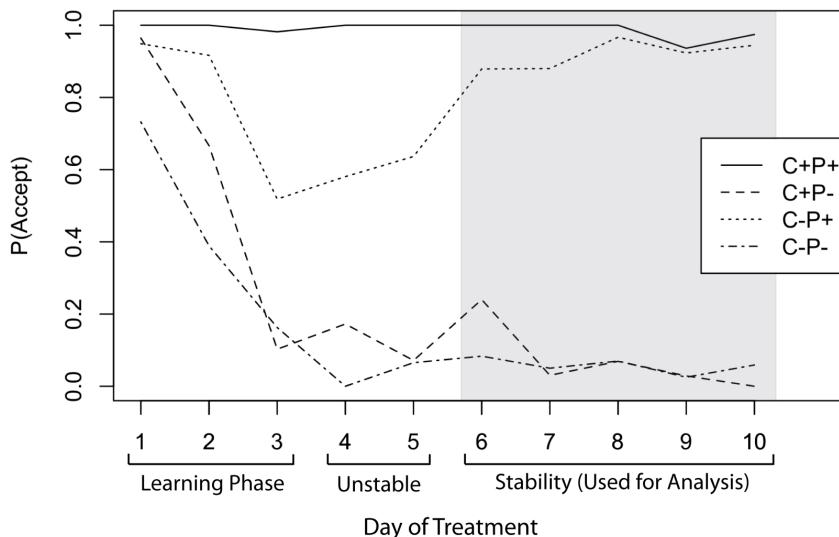


Figure 2.4: Example learning curve. Daily free trial accept rates for a single bird in a single treatment to each of the four stimulus types. Days 1-3 were the learning phase, days 4-5 were unstable, and days 6-10 showed stable responding. All data used in the analysis were collected from stable responses. This bird is showing pattern-following behavior (accepting any P+, rejecting any P-).

If behavior was still unstable after 1000 free trials had been completed, we ended the treatment anyway (this only occurred in 5 out of 72 possible cases).

2.2.7 Data Analysis

To analyze the data, we calculated the acceptance rates for each of the four possible component combinations (C+P+, C+P-, C-P+, C-P) using fourteen probe trial responses each (56 responses in total). In some cases there were more than fourteen probe trials for a given subject in a given treatment. In these cases, we randomly selected 14 probe trial responses to each of the four stimulus combinations (56 responses in total) in order to ensure that our data provided a balanced representation of the behavior of each subject in each treatment. We used these responses to calculate acceptance rates, which we then placed in a 2 x 2 table (see Table 2.1). Using these acceptance rates, we defined three dependent measures: the color effect, the pattern effect, and the interaction effect.

We calculated the color effect as the average row effect in the 2 x 2 table, that is,

		Pattern State	
		P+	P-
Color State	C+	a	b
	C-	c	d

Table 2.1: The values a , b , c , and d are acceptance rates for each of the four stimulus types, representing every possible combination of components. These rates were used to calculate the four scores used as dependent measures in the analysis.

$\frac{(a-c)+(b-d)}{2}$, following the notation shown in Table 2.1. This measures the average effect of a change in the state of the color stimulus (C+ to C-) at each of the two possible states of the pattern stimulus. We calculated the pattern effect in the same way, but using the average difference between entries in the two columns, that is, $\frac{(a-b)+(c-d)}{2}$. Both the color and pattern effect ranged from 0, indicating no effect, to 1, indicating a strong effect. We also calculated an interaction effect, which addressed whether the state of color affected the subject's behavior differently at different levels of pattern. Using the notation described in Table 2.1, the term $(a - b)$ represents the effect of pattern when color is in the C+ state, while $(c - d)$ represents the effect of pattern when color is in the C- state. We therefore used the difference $(a - b) - (c - d)$ to assess the interaction effect. In the absence of an interaction this difference should be zero. Strong interaction is indicated by values of -1 or 1, depending on the nature of the interaction. All housing and experimental procedures were approved by the University of Minnesota Institutional Animal Care and Use Committee (protocol #1109A04421).

2.3 Results

There were 8 subjects and 9 treatments, resulting in 72 subject / treatment cases. In 5 of these cases, behavior had not stabilized by the time the maximum number of trials was reached, so data was taken from a period of unstable responding. In an additional 9 cases, subjects were heavily biased toward a single button (specifically, overall acceptance rates were greater than 0.95 or less than 0.05). In the remaining 58 cases, receivers showed stable signal following behavior.

The central goal of our experiment was to test the hypothesis that receivers follow the most reliable signal components. Figure 2.5 gives an overview of the observed effects

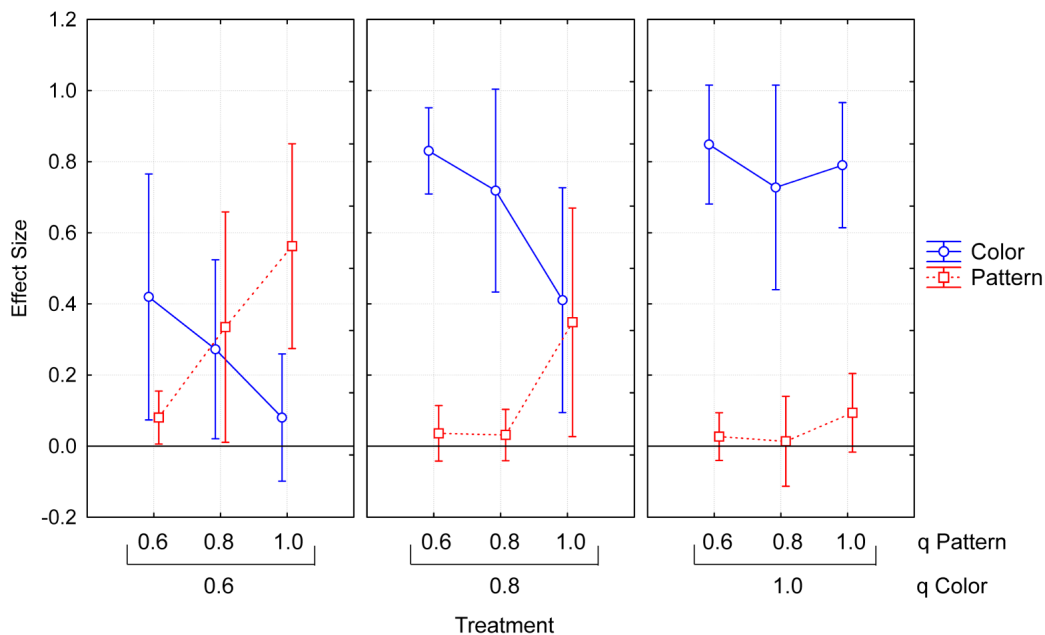


Figure 2.5: Effect of color and pattern across all treatments. Subjects strongly followed the more reliable component when reliability differences were large. Color following was seen more often than expected; all subjects followed color when it was equally reliable relative to pattern, and individual preferences were seen when color was slightly more reliable (i.e., some subjects followed color and some followed pattern). We attribute this trend to a slight color bias in our subjects (this is explored further in the Discussion section). Note that intermediate effect size values result from pooling data from color followers and pattern followers; they denote a mix of strategies across subjects rather than intermediate follow rates.

of color and pattern as a function of the experimentally manipulated differences in color reliability (q_{color}) and pattern reliability ($q_{pattern}$). We performed repeated measures ANOVA for each of these dependent measures (color effect and pattern effect), and these analyses detected no interaction between color and pattern reliability for either dependent measure (Color: $F_{4,28} = 1.248$, $p = 0.314$; Pattern: $F_{4,28} = 2.187$, $p = 0.096$). The same analyses revealed significant main effects of color and pattern reliability on both dependent measures (main effect of color reliability: Color: $F_{2,14} = 15.435$, $p < 0.001$; Pattern: $F_{2,14} = 6.728$, $p = 0.009$; and main effect of pattern reliability: Color: $F_{2,14} = 8.711$, $p = 0.003$; Pattern: $F_{2,14} = 14.330$, $p < 0.001$).

A secondary goal of our experiment was to ask whether the combinations of signal

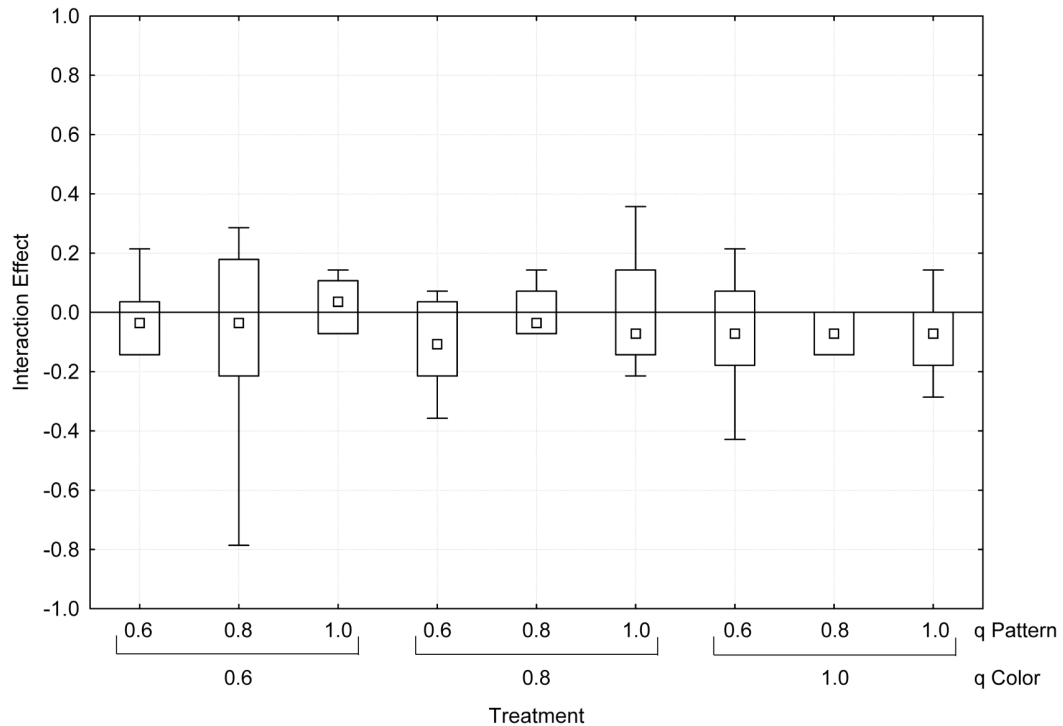


Figure 2.6: Interaction effect across all treatments. The interaction effect did not differ significantly between treatments.

components were more effective than single components. As explained in the methods, we constructed an interaction score to test this possibility. Figure 2.6 illustrates the observed interaction effect for each of the nine treatments. The mean interaction effect was near zero in every treatment (mean = -0.063 ± 0.025), and did not differ between treatments (repeated measures ANOVA: $F_{8,56} = 0.205$, $p = 0.989$).

Finally, we sought to explore the claim that receivers should only follow one stimulus component rather than some combination of both. In order to do this we classified component types (i.e., color and pattern) as either the “component of larger effect” or “the component of smaller effect.” The rationale for this approach is that, in certain treatments, we expected some birds to follow color and some to follow pattern; this approach allowed us to assess whether all subjects, as a whole, were following one or two components. Figure 2.7 shows the larger and smaller effects (white and gray bars, respectively) across all treatments. The figure shows a clear and striking difference in

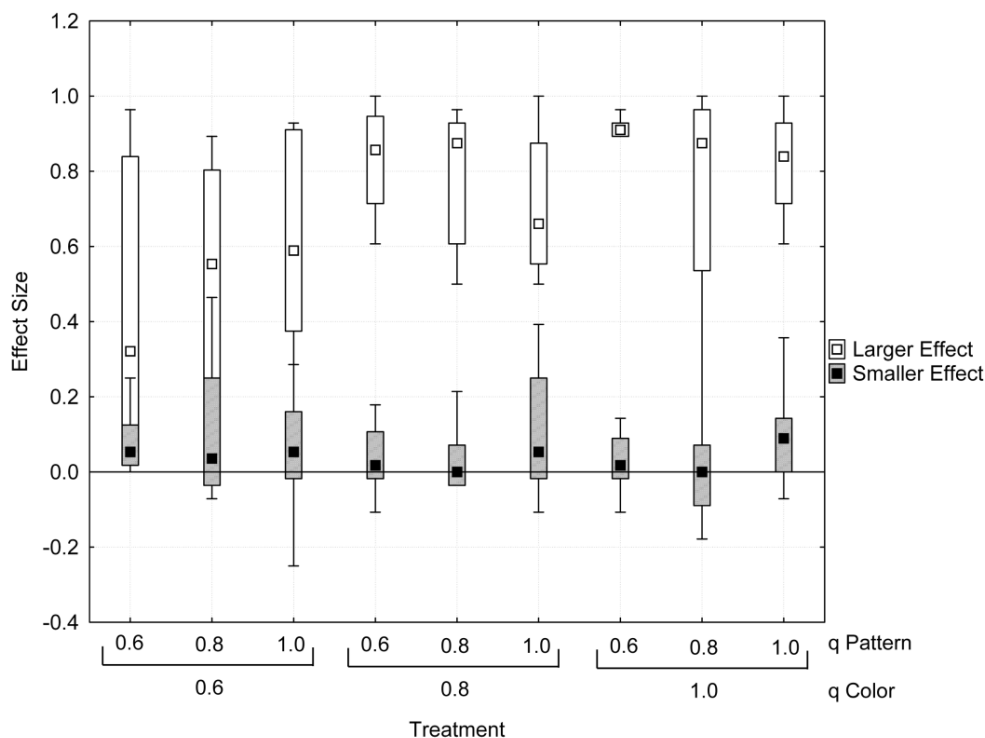


Figure 2.7: Smaller effect versus larger effect. The gray bars show the size of the effect of the “less followed” component (regardless of whether that component was color or pattern). This effect was approximately zero in every treatment and did not differ significantly between treatments. The white bars show the size of the larger effect for comparison.

the sizes of the two effects; the larger effect (mean = 0.675 ± 0.319) was roughly 11 times greater on average than the smaller effect (mean = 0.062 ± 0.133). The mean smaller effect did not differ from zero in any treatment, and did not differ significantly between treatments (repeated measures ANOVA: $F_{8,56} = 0.560$, $p = 0.777$). This result highlights an important point about the color / pattern follow rates shown in Figure 2.5. The intermediate values seen in two treatments ($q_{color} = 0.6/q_{pattern} = 0.8$ and $q_{color} = 0.8/q_{pattern} = 1.0$) do not indicate intermediate follow rates for both components. Rather, they are the result of pooling data across color- and pattern-followers.

2.4 Discussion

2.4.1 Summary of Results

Our data clearly indicate that component reliability influenced how subjects followed our experimental signals. Receivers followed color in experimental treatments where color was much more reliable than pattern and they followed pattern in experimental treatments where pattern was much more reliable than color. Another clear effect was the lack of interaction between signal components; subjects did not respond differently to different color / pattern combinations. Finally, in every case we observed a wide separation between the rates at which the two components were followed. One component was followed at a high rate, while the other was followed at a very low rate or ignored completely.

Despite these strong trends, the results differed from our predictions in two notable ways. First, our data suggest that the secondary component may have a small main effect. Though our analysis did not find the “smaller effect” to be significant, Figure 2.7 illustrates that the values are generally positive, rather than equally distributed about zero as might be expected. We were not able to confirm this effect in this study, however, this could be an interesting phenomenon to explore; perhaps receivers utilize “extra” components in a manner distinct from how they use the primary component (e.g., following at a low background rate). Second, we observed a bias for following color. While the predictions of Figure 2.1 held up nicely in the extreme treatments (the upper left and lower right corners), the region of color-following was larger than expected (the midline was shifted down). That is, all subjects followed color when the two components were equally reliable, and we observed a mix of strategies (i.e., some subjects followed color and some followed pattern) when color was somewhat less reliable than pattern. This suggests that our subjects preferred to follow color, and occasionally did so to the detriment of their performance.

These results broadly support the predictions of several economic models of receiver behavior, which indicate that following the single most reliable signal component is typically the best strategy (e.g., Bro-Jørgensen 2010; Iwasa and Pomiankowski 1994; Johnstone 1995; Pomiankowski and Iwasa 1993; Schluter and Price 1993; Wilson et al. 2013). However, we also observed a notable bias for following color. The color bias in

equal-reliability treatments was not inconsistent with our model; however, some subjects also followed color when pattern was somewhat more reliable, which is not an economically valid strategy. These results suggest that something other than component reliability made color a more salient stimulus in general. We discuss this idea further in the next section.

2.4.2 Connections to Salience

This study is related to the psychological concept of stimulus salience in several ways. The term “salience” is used in many disciplines in the behavioral sciences to describe situations in which one type of experience is more prominent or influential than another. In this broadest sense, one can interpret our experiment as a study of the effects of component reliability on component salience. However, investigators also use the term salience in a narrower sense that stems from Rescorla and Wagner’s (1972) influential model of classical conditioning. Rescorla and Wagner conceived of salience as a fixed stimulus property determined by the physical attributes of the stimulus (brightness, contrast, etc.) and the animal’s sensory apparatus. This interpretation of salience is popular with behavioral biologists (e.g., Kazemi et al. 2014), presumably because it complements ideas from the biological signaling literature such as sensory drive, sensory exploitation, and pre-existing sensory biases (e.g., Endler and Basolo 1998; Ryan and Cummings 2013). In this narrower “fixed property” sense, our study could not be seen as a study of salience because the differences in component effects we report are learned differences and not fixed properties of the underlying stimuli. We point out, however, that psychologists began chipping away at the fixed property interpretation of stimulus salience almost immediately after the publication of the Rescorla-Wagner model (see Mackintosh 1973; Pearce and Bouton 2001; Pearce and Hall 1980). For example, in the conditioning phenomenon of latent inhibition, investigators find that prior exposure to a stimulus in the absence of reinforcement contingencies (say a green dot that is always in the background predicting nothing) inhibits later conditioning with this stimulus. Rescorla himself suggested that one could interpret this as learned change in stimulus salience (Rescorla and Wagner 1972, p. 30). Rescorla’s thinking here is very much in line with the hypothesis tested in this paper.

From the broadest biological perspective, it would seem reasonable to think of

salience as an attribute of the behavioral phenotype that, like nearly every other aspect of behavior, is controlled by a complex interaction between genetically determined effects and experientially controlled effects. If we accept this premise, then this experiment can be seen as a study of one experientially controlled attribute of stimulus salience (component reliability). However, the study also shows the importance of pre-existing salience differences in the sense that it shows a bias favoring the color component of our experimental stimuli. Studies suggest that birds have innate color biases (e.g., Cook and Roper 1989; Schmidt and Schaefer 2004), as well as an innate tendency to weigh color differences more heavily than other types of cues (Aronsson and Gamberale-Stille 2008, 2012).

If narrow-sense salience is an evolved feature of behavior, how would natural selection act on it? One reasonable hypothesis parallels the model developed here. Selection should act to increase the salience of stimuli that reliably predict biologically meaningful outcomes, and decrease the salience of unreliable stimuli. This hypothesis is logically identical to our hypothesis, except that the reliability in question is reliability experienced over the selective history of the focal species. Dunlap and Stephens (2014) have recently tested exactly this hypothesis using the techniques of experimental evolution in replicate populations of *Drosophila*. Their results parallel the results presented here by showing that reliability experienced across selective history seems to shape biases, just as this study shows that reliability experienced within an animal's lifetime influences the weight that signal component receives.

2.4.3 Alternative Explanations for the Prevalence of Complex Signals

In this study, receivers rarely, if ever, followed multiple components. However, we do not suggest that complex signals are never beneficial for the receiver; indeed, there is extensive experimental evidence that multiple signal components can be important (e.g., Kelly and Marples 2004; Leonard et al. 2011; Siddall and Marples 2008; Uetz et al. 2009; VanderSal and Hebets 2007). Researchers have proposed a variety of specific conditions that might favor complex signal following by receivers (see Wilson et al. 2013 for a recent overview). A particularly compelling argument is that complex signals (particularly multimodal signals) might be more detectable in environmental noise (Candolin

2003; Hebets and Papaj 2005), which is a ubiquitous problem in natural signaling systems. Another popular hypothesis, the multiple messages hypothesis, states that signal components might not be truly redundant; that is, different components may reflect different aspects of an underlying quality (Hebets and Papaj 2005; Johnstone 1995; Moller and Pomiankowski 1993). Several authors have argued that complex signal following is economically stable under a multiple messages framework (Johnstone 1995; Moller and Pomiankowski 1993; Wilson et al. 2013).

From the perspective of a purely strategic design, complex signals should be unstable; a single discriminable difference should be sufficient to signal an underlying state to a receiver (Wilson et al. 2013). This prediction makes two basic assumptions that are supported across a range of signaling systems; namely, that signals are reliable (e.g., Alonso-Alvarez et al. 2004; Ballentine et al. 2008; Blount et al. 2009; Kilner 1997), and that receivers are sensitive to signal reliability and alter their behavior accordingly (e.g., Blumstein et al. 2004; Gamberale-Stille and Guilford 2004; Skelhorn and Rowe 2006). However, unreliable cues can also influence receiver behavior in certain contexts (e.g., Endler and Basolo 1998; Moller and Pomiankowski 1993; Rendall et al. 2009). Multi-component signals can pair reliable and unreliable cues, and researchers have proposed several ways in which following such signals might be a stable receiver strategy. Alerting and attention-altering hypotheses predict that a poorly reliable but conspicuous signal component might function to draw receiver attention to a second, reliable component (reviewed in Hebets and Papaj 2005). Rowe (1999) reviews several other ways in which a secondary signal component can improve the detectability, discriminability, and memorability of a primary component. Sexual signaling models have indicated that preferences for ornaments or Fisherian traits may be able to persist alongside preferences for reliable traits, provided that preference for the unreliable trait is weak (Iwasa and Pomiankowski 1994) and costs are low (Pomiankowski and Iwasa 1993). Moller and Pomiankowski (1993) hypothesized that this would be a dynamic process; that is, preferences for individual unreliable traits might not be stable, but over time the tendency to follow multiple components could be. The role of reliability in signaling is somewhat contentious, with several authors arguing that other considerations, such as transmission and receiver processing (Guilford and Dawkins 1991, 1993) or the ability to exploit receiver biases (Endler and Basolo 1998; Owren et al. 2010; Rendall et al. 2009) may

play equal or greater roles in signal design. Complex signals could conceivably combine components optimized for different signaling roles, adding a new element to this debate.

2.4.4 Concluding Summary

Economic models of receiver behavior indicate that, under basic conditions, following multiple redundant signal components is not a stable receiver strategy (e.g., Bro-Jørgensen 2010; Iwasa and Pomiankowski 1994; Johnstone 1995; Pomiankowski and Iwasa 1993; Schluter and Price 1993; Wilson et al. 2013). To our knowledge, this study has provided the first experimental test of this prediction. Our findings largely agree with these models; receivers followed a single component at a high rate, and either ignored the second component or followed it at a very low rate. Component reliability had a clear and consistent effect on the extent to which receivers followed a signal component, although we also found that receivers tended to “over follow” the color component of our experimental signal.

Chapter 3

Does multimodality *per se* improve receiver performance? An explicit comparison of multimodal versus unimodal complex signals¹

3.1 Introduction

Signals with multiple components, known as complex signals, have become a popular topic in the field of animal communication over the past decade. Increasingly, researchers are focusing on multimodal signals, defined as signals with components that occur across multiple sensory modalities – for example, a courtship signal that includes a dance and a vocalization, or a fertility signal that incorporates a color and an odor. Multimodal signals are prevalent in animal communication, and several authors have argued that multimodal signals are more effective than signals confined to a single sensory modality

¹ This chapter is currently in revision for *Behavioral Ecology and Sociobiology* as Rubi, T.L. and Stephens, D.W. Does multimodality *per se* improve receiver performance? An explicit comparison of multimodal versus unimodal complex signals in a learned signal following task.

(e.g., Kulahci et al. 2008; Partan 2013; Rowe and Halpin 2013; Siddall and Marples 2008). It is not yet clear whether multimodal signals and unimodal complex signals, defined as signals with multiple components in the same modality, should be treated as fundamentally different; indeed, this is the first of Partan’s (2013) top ten unanswered questions in multimodal communication.

A major proposed benefit of multimodal signals is that they improve the performance of receivers (e.g., Kulahci et al. 2008; Partan and Marler 2005; Rowe and Halpin 2013; Siddall and Marples 2008; VanderSal and Hebets 2007), and many authors have argued that multimodal signals may be processed more efficiently simply by virtue of the fact that they stimulate multiple sensory systems (Guilford and Dawkins 1991; Hebets and Papaj 2005; Rowe 1999; Siddall and Marples 2008). Multimodal signals are likely processed differently than unimodal signals, however, we still have much to learn about the mechanisms of multisensory versus unisensory neural processing (see Alvarado et al. 2007; Gingras et al. 2009; Pluta et al. 2011 for work in this area). At the behavioral level, many empirical studies have found that receivers exhibit enhanced responses to multimodal signals relative to isolated components (e.g., Hughes 1996; Narins and Grabul 2005; Siddall and Marples 2008; Stynoski and Noble 2011; Uetz et al. 2009; VanderSal and Hebets 2007), however, few studies have directly compared multimodal signals to unimodal complex signals (though see Hauglund et al. 2006; Taylor and Ryan 2013 for relevant work). Comparing whole multimodal signals to isolated components confounds the effect of multiple modalities with the effect of multiple components. This approach is problematic for two reasons. First, when testing responses to natural signals (as most of these studies do), reduced responses to isolated components may simply be a result of testing “partial signals” rather than a result of modality effects *per se*. Second, complex signals themselves have been hypothesized to improve receiver performance (e.g., Hebets and Papaj 2005). To truly test for modality effects, the signals being compared should have the same number of components. Our aim in this study was to directly compare how receivers used two-component signals that differed only in whether they were cross-modal.

In this experiment, we focused on learned signal following using artificial signals and captive blue jays (*Cyanocitta cristata*) as our study subjects. Inspired by aposematic signaling, our experimental signals were computer-generated butterfly images. Each

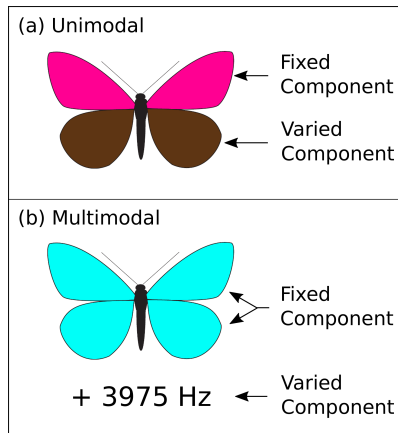


Figure 3.1: Example stimuli. We refer to component 1 as the fixed component, since it is a color in both unimodal and multimodal treatments. We refer to component 2 as the varied component, since it is a color in unimodal treatments and a tone in multimodal treatments. (a) A unimodal stimulus: component 1 is the forewing color and component 2 is the hindwing color. (b) A multimodal stimulus: component 1 is the color of the wings and component 2 is a pure tone.

signal had two components that were independently informative about the true state of our virtual butterfly (“good” or “bad”). Subjects were rewarded for accepting butterflies in the good state and rejecting butterflies in the bad state. The first component was a visual component (color) in unimodal and multimodal treatments – we refer to this component as the fixed component. The second component was a visual component (color) in unimodal treatments, and an acoustic component (a pure tone) in multimodal treatments – we refer to this component as the varied component (see Figure 3.1 for example stimuli). Each component could be in one of two states; it could either indicate that the true state was good or that the true state was bad. If the component was color, the hue of the color (e.g., red vs. green) was the informative difference; if the component was sound, the frequency of the tone (e.g., 1000 Hz vs. 2000 Hz) was the informative difference.

We also experimentally varied the reliability of signal components. Signal reliability is the degree to which the signal state (or component state) accurately reflects the true state. We quantified reliability with the variable q , which we defined as the conditional probability: $q = P(\text{true state is good} \mid \text{component state indicates “good”}) = P(\text{true$

state is bad | component state indicates “bad”). We tested two levels of reliability (0.7 and 0.95) in a full factorial design, resulting in four reliability pairings. At each reliability pairing we completed two treatments, a unimodal treatment and a multimodal treatment, resulting in a grand total of eight treatments. We then characterized how multimodality and reliability influenced general performance, component preference, and integration of the two components.

3.2 Methods

All housing and experimental procedures were approved by the University of Minnesota Institutional Animal Care and Use Committee (protocol #1109A04421).

3.2.1 Subjects

Our subjects were eight adult blue jays of unknown sex, randomly selected from a captive colony. We housed subjects in the testing chambers for 23 hours per day; they were removed for an hour for health checks, cleaning, and maintenance. The experimental day began each morning at 07:00 and ran until 15:00. Subjects received all of their food from the apparatus except when weight dropped dangerously low. We provided ad libitum water and maintained subjects on a 12 hour light:dark cycle.

3.2.2 Testing Apparatus

We built each testing apparatus into a separate sound attenuation chamber (Figure 3.2). We presented visual stimuli on an LCD screen at the front of the box, and acoustic stimuli from a speaker attached to the chamber lid. The food cup was located directly under the screen, and the two response keys (MED Associates ENV-123AM) were located on either side. Reject keys were marked with an X and Accept keys were unmarked, and the positions of the two keys were randomized across subjects. A response perch and indicator light were located at the rear of the box. A MED-PC computer program (MED Associates, Burlington, VT, USA) controlled inputs and outputs and recorded data.

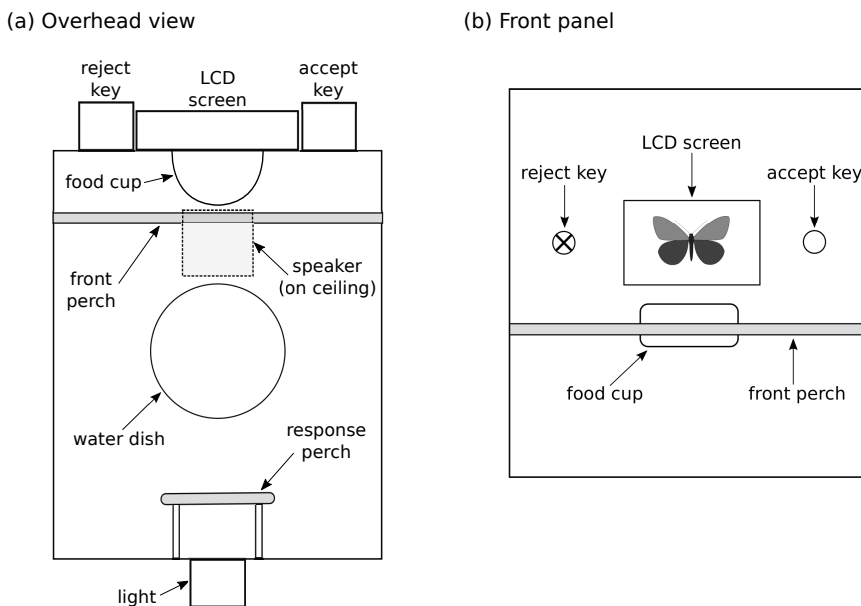


Figure 3.2: An example configuration of the testing apparatus (the left-right positions of the accept and reject keys were randomized across subjects) (a) Overhead view. (b) Front panel view.

3.2.3 Stimuli and Treatment Design

We generated the visual stimuli using Adobe Illustrator (Adobe Systems, San Jose, California) and the acoustic stimuli using Audacity (Audacity Team). In unimodal treatments, stimuli were butterfly stencils with two colors corresponding to the forewing and the hindwing (Figure 3.1a). We designed the stencils such that the areas taken up by the two colors were equal in size. In both unimodal and multimodal treatments, all colors used within a single treatment were spaced at least 80 degrees apart on the Adobe Illustrator color wheel. In multimodal treatments, the butterfly stimuli had a single color on both wings and were always accompanied by a pure tone (Figure 3.1b). The eight tones used for multimodal treatments were evenly spaced from 1000 Hz to 3975 Hz, within the range of peak sound sensitivity for blue jays (Cohen et al. 1978). The tones paired together in a treatment differed by at least 850 Hz. Tone amplitude at the rear perch during playback was at least 60 dB.

We constructed eight stimulus sets by pseudo-randomly combining colors and tones (following the restrictions just described). These eight stimulus sets were used for

all subjects. Each subject saw each stimulus set in only one treatment. We pseudo-randomly assigned stimulus sets to treatments such that two subjects saw each stimulus set in a given treatment. For each subject, we randomly assigned the colors and tones as the positive and negative components (that is, indicating “good” and “bad,” respectively). This randomization scheme controlled for effects of treatment order, stimulus set, and individual colors and tones.

3.2.4 Trial Walkthrough

Trials were separated by an intertrial interval (ITI) of 110 seconds. At the start of each trial, the computer determined the trial type (free or forced) and the true state (good or bad) based on the treatment parameters. The computer then independently assigned the states of the signal components probabilistically based on the component reliabilities. To begin the trial, a flashing light at the rear perch indicated that the trial was ready to begin. The subject initiated the trial by hopping to the rear perch, which caused the butterfly stimulus to appear on the screen and, if applicable, the acoustic playback to begin. The two response keys were also illuminated. The signal was presented continuously until the subject responded; that is, the visual stimulus remained displayed on the screen and acoustic components played continuously. Subjects pecked the accept key or the reject key to respond, which immediately extinguished the response keys and ended the signal playback. Correct responses were rewarded with two food pellets, accompanied by a flashing magazine light. When subjects responded incorrectly, the signal and keys extinguished and no food was delivered. If a subject had not responded within 15 minutes, the trial aborted, the ITI restarted, and the trial was repeated.

3.2.5 Trial Types

We arranged trials in blocks of 40, beginning with 8 forced trials followed by 32 free trials. Forced trials (also known as “no choice” trials) are similar to free trials, except that they require subjects to respond in a predetermined way and experience the associated reward; this ensures that subjects experience all possible choices and outcomes in a treatment. For example, in a forced correct accept trial, the computer assigned the component states and the apparatus illuminated the signal and accept / reject keys,

exactly as in free trials. However, the trial only ended when the subject pecked the accept key (that is, the reject key was not responsive). The subject received a reward in forced correct trials, but not forced incorrect trials. Subjects responded freely in free trials and were rewarded according to their responses.

3.2.6 Stability Assessment

Subjects continued in a treatment until signal following behavior had stabilized, and responses from the period of stable responding were used in the analysis. The first 600 trials of each treatment constituted the learning period – trials from the learning period were not used to assess stability. After the learning period had ended, the treatment continued until the subject completed three consecutive days of stable signal following (hereafter referred to as the stability period). Stable signal following was defined as follows: 1) the overall accept rate for all pooled stimuli was greater than 0.1 and less than 0.9, and 2) the accept rates for each stimulus differed by less than 17% across the three days. The first criterion ensured that subjects weren't simply accepting or rejecting every stimulus, and the second criterion ensured that responses to each stimulus had stabilized. If these criteria were not met by the time subjects had completed 3,500 trials, the treatment ended and the final three days were used in the analysis.

3.2.7 Data Analysis

All subjects completed all treatments in a random order, and we analyzed the data using a within-subjects design. For all dependent measures except learning rate, we focused on data from the stability period (or, if responses never stabilized, the last three days before the cutoff of 3,500 trials was reached). All free trials from the stability period were included in the analysis – typically around 550 trials per subject per treatment.

We assessed performance based on three measures: the proportion of correct responses ($P(\text{Correct})$), the reaction time, and the learning rate. The reaction time was measured as the time elapsed between the presentation of the signal on the screen and the moment the subject registered a response. Learning rate was defined as the total

number of trials completed in a treatment – in other words, the total number of trials completed before the stability criterion was met. If behavior had not stabilized by the end of the treatment, the total number of trials was set to 3,500 (the maximum treatment length).

We next assessed which component influenced decisions more strongly. To do this, we examined responses to the subset of signals in which the components disagreed – that is, one component indicated that the true state was good and the other indicated that the true state was bad. We focused on these cases because responses to the signals in which components agreed did not provide information about preference. We then determined the proportion of “follow responses” for the varied component ($P(\text{Follow Varied Component})$). We defined a follow response as accepting a component signaling “good” or rejecting a component signaling “bad.” Recall that in unimodal treatments the varied component was a color, and in multimodal treatments it was a sound – thus, this measure allowed us to test for biases toward sound or color cues. The follow rate for the fixed component was necessarily the inverse of $P(\text{Follow Varied Component})$ because following the fixed component and following the varied component were mutually exclusive strategies for disagreeing signals. Therefore, we report the statistics for following the varied component only.

Finally, we assessed whether receivers were more likely to integrate both components into decision-making when signals were multimodal. To quantify this measure, we used the responses to all four component combinations to calculate a value that we called the interaction score. This measure quantified the extent to which receivers responded differently to different combinations of components. Following the notation described in Table 3.1, the term $(a - b)$ measures the effect of the varied component when the fixed component is in the “signal good” state, while $(c - d)$ measures the effect of the varied component when the fixed component is in the “signal bad” state. We therefore use the difference $(a - b) - (c - d)$ to assess the interaction effect. In the absence of an interaction (i.e., if subjects only follow one component) this difference should be zero. Strong interaction is indicated by values of -1 or 1, depending on the nature of the interaction.

To summarize, we calculated five dependent measures: 1) $P(\text{Correct})$, 2) reaction time, 3) learning rate, 4) $P(\text{Follow Varied Component})$, and 5) interaction score. We

		Varied Component State	
		“Signal Good”	“Signal Bad”
Fixed Component State	“Signal Good”	<i>a</i>	<i>b</i>
	“Signal Bad”	<i>c</i>	<i>d</i>

Table 3.1: Accept rates to all four possible combinations of components. The values *a*, *b*, *c*, and *d* are acceptance rates. These values were used to calculate the interaction score to assess component integration.

analyzed each of these dependent measures using a $2 \times 2 \times 2$ repeated measures analysis of variance with factors of modality (unimodal or multimodal), reliability of the fixed component (0.7 or 0.95), and reliability of the varied component (0.7 or 0.95).

3.3 Results

3.3.1 Performance Measures

We found no significant difference in the proportion of correct responses during stable signal following in multimodal versus unimodal treatments (Figure 3.3). A repeated measures ANOVA indicated a significant interaction between the reliability of the fixed component and the reliability of the varied component (repeated measures ANOVA: $F_{1,7} = 52.008$, $p < 0.001$), however, there were no significant interactions involving modality, nor a significant main effect of modality. We also found no significant difference in the learning rate in multimodal versus unimodal treatments, measured by the number of trials completed to achieve the stability criterion (Figure 3.4). A repeated measures ANOVA indicated a significant effect of the reliability of the varied component only ($F_{1,7} = 7.066$, $p = 0.033$). There were no significant effects of modality or any interactions involving modality. Finally, we assessed the response time during stable responding. The average response time was 30.39 ± 0.8 tenths of a second. The response time data violated the sphericity assumption, so we applied a square root transformation. We found no significant differences in response time (square root transformed) due to reliability, modality, or any interactions. We performed a power analysis following Cohen’s (1988) framework and determined that we had sufficient power to detect moderate effects with our design.

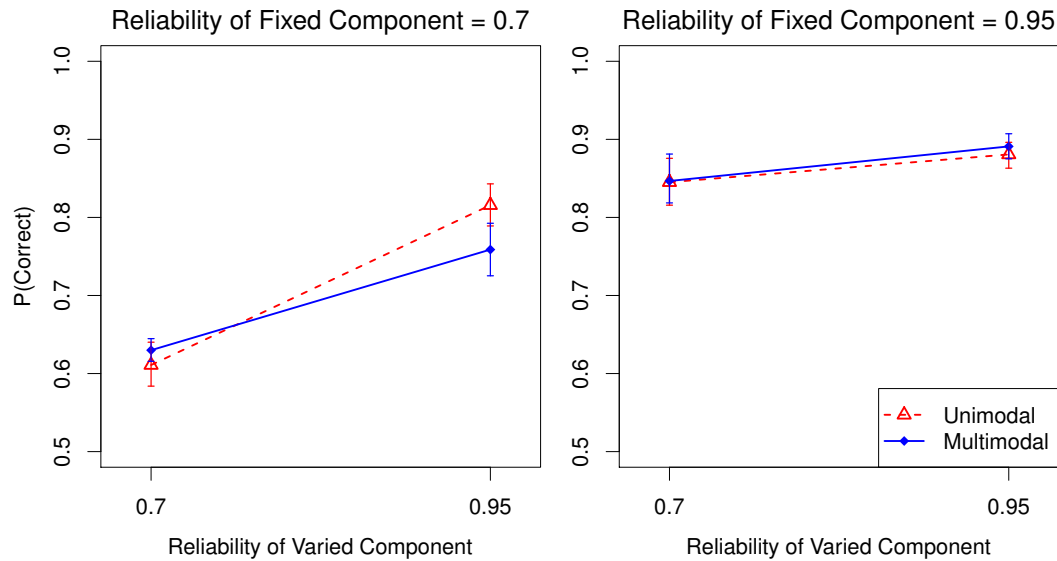


Figure 3.3: Proportion of correct responses across all treatments. A repeated measures ANOVA revealed a significant interaction between the reliability of the fixed component and the reliability of the varied component. There were no significant effects of modality or interactions involving modality.

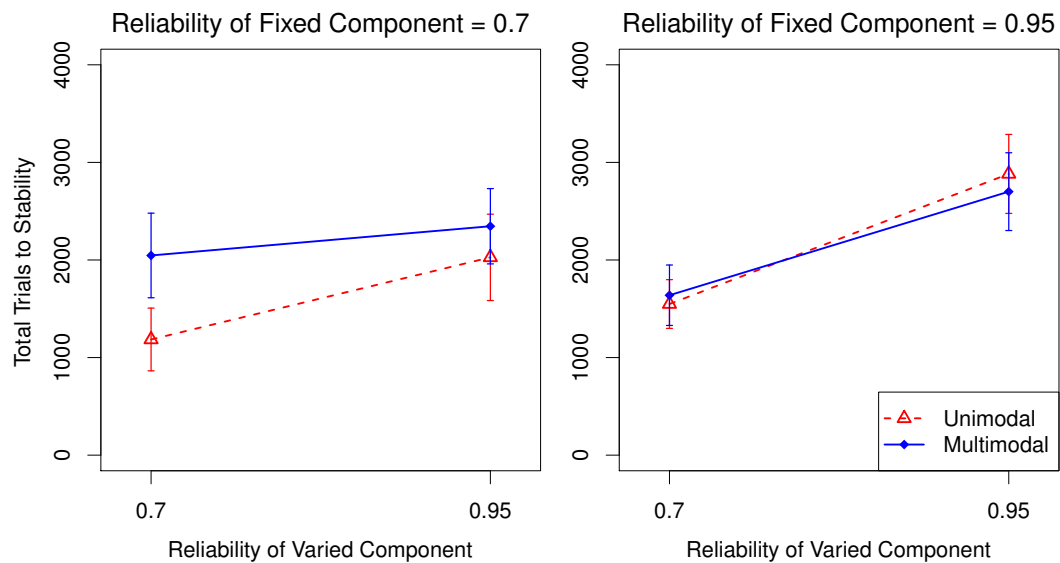


Figure 3.4: Number of trials until stability criterion was reached. A repeated measures ANOVA revealed a significant main effect of the reliability of the varied component. There were no significant effects of modality or interactions involving modality.

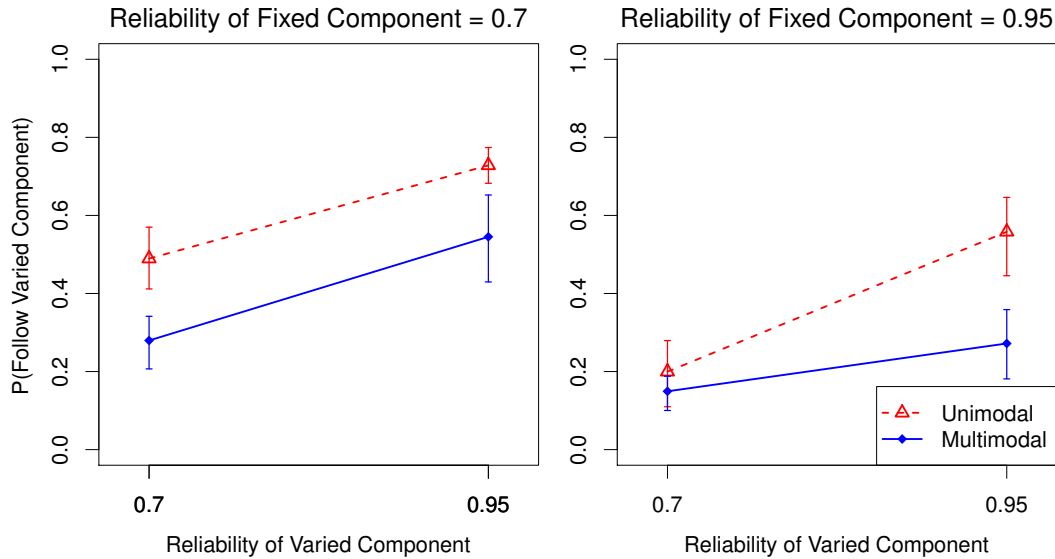


Figure 3.5: Proportion of “Follow Varied Component” responses when the two stimuli disagreed. A repeated measures ANOVA revealed significant main effects of modality, the reliability of the fixed component, and the reliability of the varied component.

3.3.2 Component Preference

Most subjects followed sound at a higher rate than color in at least one treatment, however, subjects had a tendency to prioritize color cues. This trend was most apparent in the two treatments where sound and color were equally reliable; in these treatments, seven out of eight subjects followed color (the fixed component) at a higher rate than sound (the varied component). In the analogous unimodal treatments, preferences for following the forewing color (the fixed component) or hindwing color (the varied component) were much more evenly split. Two subjects did not follow sound at a higher rate than color in any treatment. A repeated measures ANOVA revealed that $P(\text{Follow Varied Component})$ was significantly influenced by modality ($F_{1,7} = 13.387$, $p=0.008$), as well as the reliabilities of the fixed component ($F_{1,7} = 31.827$, $p < 0.001$) and the varied component ($F_{1,7} = 14.851$, $p=0.006$). The interaction between modality and the reliability of the varied component bordered on significance ($F_{1,7} = 3.995$, $p=0.086$) (Figure 3.5).

3.3.3 Component Integration

The interaction score hovered around zero in every treatment (Figure 3.6). A repeated measures ANOVA no significant differences in the interaction score as a result of modality, reliability, or any interactions.

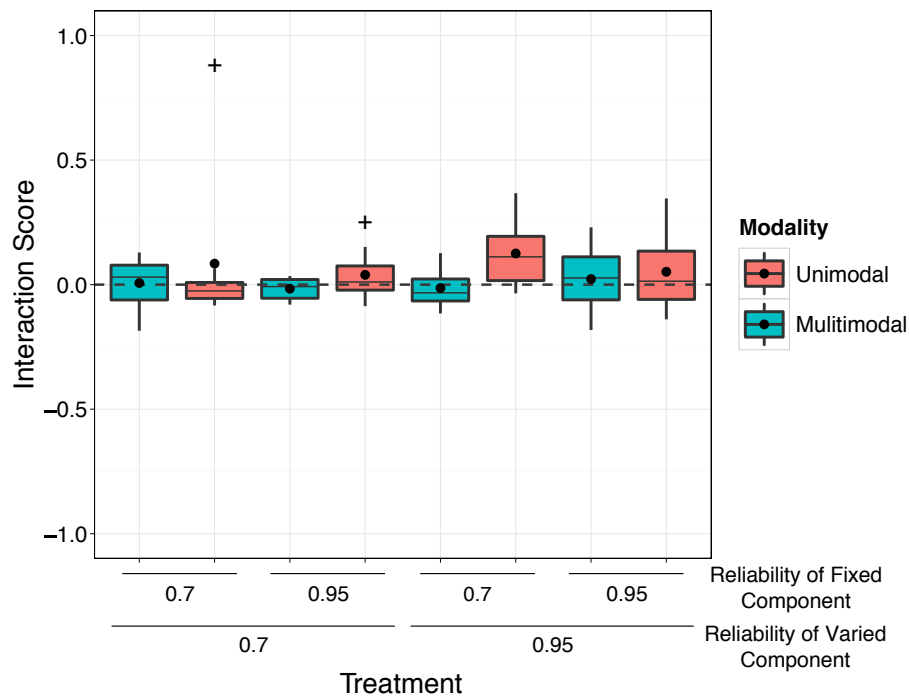


Figure 3.6: Interaction scores across all treatments. The black dots indicate means, lines indicate medians, box edges span the first to third quartiles, whiskers extend to 1.5 * interquartile range, + indicates an outlier.

3.4 Discussion

3.4.1 Summary of Results

As expected, receivers were sensitive to component reliability; the reliabilities of both components significantly influenced the proportion of correct responses during stable responding, as well as the number of trials completed before stability was reached. Most

receivers readily followed both color and sound components, however, we found no evidence that our experimental receivers learned multimodal signals any differently than they learned complex unimodal signals. None of the three performance measures we examined (proportion of correct responses, learning rate, and response time) differed significantly between unimodal and multimodal treatments. We also found no evidence that receivers were more likely to integrate multiple components into decision-making when signals were multimodal. Indeed, consistent with our findings in Chapter 2, we found little evidence that receivers ever integrated multiple components; instead, they followed one component at a high rate and the other at a minimal rate (if at all). In other words, we found no evidence that multimodality per se was beneficial for receivers in this study. These results certainly do not rule out the hypothesis that multimodality improves receiver performance, however, they call into question the generality of such an effect. How then do we explain the prevalence of multimodal signals in natural communication systems? We argue that the most likely explanation is that multimodal signals can be broadly beneficial, but these benefits are contingent upon other conditions. We briefly explore this idea below.

We also observed a notable sensory bias in our subjects; subjects tended to prioritize color cues over sound cues. The observed color preference in our subjects makes sense given avian sensory ecology. Birds have excellent color vision, and likely use visual cues more than acoustic cues when foraging and assessing aposematic signals (though acoustic aposematism is seen in several insect groups with avian predators - e.g., Brown et al. 2007; Bura et al. 2009; Masters 1979). While an interesting side note, we do not believe that this bias affects our core results. Most birds followed sound at a higher rate than color in the treatment in which sound was the more reliable component. Importantly, we did not see decrements in performance in that treatment – if we had, we would have observed significant interactions between modality and the component reliabilities when analyzing the performance measures. Similarly, we saw no difference in interaction scores, indicating that sound-following birds did not incorporate color components into decision-making despite the bias toward following color cues. This illustrates that while our subjects had a tendency to prioritize color cues, this bias was easily overcome. Furthermore, such minor biases are common across taxa and should not preclude the use of multimodal signals by receivers.

It is important to be explicit about the conclusions we can draw from this experiment, as well as the extrapolations we can make to other systems. In this study, we examined learned signal following behavior to novel stimuli in modalities toward which our study species had minimal prior biases; we believe that our results would likely be replicated in analogous studies focusing on different taxa. However, it would be erroneous to conclude from our study that, for example, no receiver species will show enhanced responses to visual-acoustic signals, or that blue jays will not show enhanced responses to visual-olfactory signals. It is not yet clear how different types of multisensory stimuli are processed, or how such processes differ across taxa. It seems likely that some receivers will possess special adaptations to process certain categories of multimodal signals (in a sense, such receivers could be thought of as multisensory specialists, analogous to sensory specialists). Heberts and Papaj (2005) outline a scenario by which these adaptations might arise in a separate context and subsequently be co-opted for communication, however, it also seems plausible that such adaptations could first arise in a communicative context. Studies such as this one, which focus on receivers lacking such innate enhanced responses, could help distinguish between these hypotheses. This work can complement studies examining system-specific biases to create a broader framework for understanding how multimodal signals, and the preferences for them, arise.

3.4.2 Alternative Benefits of Multimodal Signals

While our results appear to contradict the growing consensus on the benefits of multimodality, they are actually in agreement with economic predictions of stable signal use. For example, in a recent paper Wilson et al. (2013) comprehensively illustrated that, in the absence of further assumptions, it is not clear that a basic multimodal signal (or indeed, any complex signal) offers any advantage over a single-component signal. This experiment supports this assertion (also see Rubi and Stephens 2015). Nevertheless, multimodal signals are widespread in animal communication. Wilson et al. (2013) also outline a number of constraints on signaling systems that can make following multimodal signals a beneficial receiver strategy. We argue that these additional constraints may explain why multimodality is beneficial in many (and perhaps most) signaling systems.

A particularly compelling argument is that multimodal signals are better able to

overcome environmental noise (Candolin 2003; Hebets and Papaj 2005; Wilson et al. 2013). For example, a vocalization is relatively unhindered by dense foliage, while bright plumage is unaffected by acoustic background noise. Environmental noise is a ubiquitous problem in natural signaling systems – it is conceivable that this benefit alone could explain the prevalence of multimodal signals in nature. Some other possibilities are that signal components in different modalities provide enhanced attentional benefits (Hebets and Papaj 2005) or target different types of receivers, particularly in aposematic signals with multiple receiver species (Andersson et al. 2002; Ratcliffe and Nydam 2008; Wilson et al. 2013). A comprehensive overview of constraints that could favor multimodal signals is outside the scope of this paper, however, for more information we refer readers to the review in Wilson et al. (2013). Future work can build on our approach to address such hypotheses, for example, by manipulating noise levels, or by testing behavior in non-cued trials (which may enhance attentional effects).

3.4.3 Concluding Summary

With this experiment, we have attempted to take the first step in addressing the question of whether unimodal complex signals and multimodal signals are fundamentally different. We found that, given a two-component, redundant, reliable signal in a noise-free environment, receivers performed equally well when signals were unimodal or multimodal. This suggests that multimodal signals do not provide basic informational benefits for receivers, in agreement with previous economic predictions (e.g., Wilson et al. 2013). The next step is to systematically test receiver behavior under constrained conditions that have been predicted to favor multimodal signal following. Given the prevalence of multimodal signals in nature, such studies may have much to tell us about animal communication.

Chapter 4

An experimental test of the “Confirm the Rare State” hypothesis

4.1 Introduction

Increasingly in the field of animal communication, signals are being recognized as complex displays incorporating many different components – colors, patterns, sounds, odors, etc. A commonly cited benefit of such signals (called “complex signals”) is that they improve the performance of receivers. Researchers have identified a number of ways that receivers can benefit from complex signals. For example, multiple signal components can facilitate signal detection and learning (Candolin 2003; Hebets and Papaj 2005; Rowe 1999) or improve performance in environmental noise (Candolin 2003; Wilson et al. 2013). Despite these successes, several researchers have commented on a persistent, somewhat unexpected result: from a purely informational perspective, it is unclear why signals with multiple components should be any better at conveying a message than a single signal is (Bro-Jørgensen and Dabelsteen 2008; Johnstone 1995; Rubi and Stephens 2015; Schluter and Price 1993; Wilson et al. 2013). For a redundant, reliable signal, in the absence of further constraints, receivers will do as well or better following a single signal component than they will by following multiple components.

Given that complex signals are practically ubiquitous in animal communication, this observation is something of a puzzle. In this paper, we argue that complex signals may actually be economically advantageous more often than previously recognized; namely, when one signaler state (e.g., good male, toxic prey) is rare. We developed this idea and the corresponding model in a previous work (Rubi and Stephens *in revision*). Here, we reiterate the key points and present results from the first experimental test of our model. We refer readers to Rubi and Stephens (*in revision*) for the full derivation of the model.

We focus on a simple case of signal following. A receiver is presented with a signal (e.g., a mating display), and must choose between two actions, which we call accept and reject (e.g., mate and don't mate, respectively). We assume that signalers can be in one of two true states, which we call good and bad (e.g., high quality male and low quality male). A receiver benefits from accepting a signaler in the good state and rejecting a signaler in the bad state. A signal component (e.g., a color patch) can also be in two states. The component can indicate that the true state is good (e.g., bright red patch) or indicate that the true state is bad (e.g., dull red patch). We refer to the signal component as S , and the two states as $S+$ (indicates true state is good) and $S-$ (indicates true state is bad). When S is completely reliable, the component state ($S+$ or $S-$) always correctly indicates the true state. In this case, a receiver always benefits from following component S ; she observes the component state (e.g., $S+$), performs the matching action (accept), and benefits because her behavior matches the true state (she accepts a good signaler). Components are not always completely reliable, however. We define a component's reliability as the probability that the component state matches the true state. We refer to the reliability of S as q , and define it as the conditional probability: $q = P(S+ | \text{Good}) = P(S- | \text{Bad})$. Now, we simply add another component, T (e.g., a vocalization). T can also be in two states, $T+$ (e.g., low frequency) and $T-$ (e.g., high frequency), and these states are informative about the same true state. We refer to the reliability of T as r , and define it in the same way as q : that is, $r = P(T+ | \text{Good}) = P(T- | \text{Bad})$.

We assume that S and T are independent of each other (that is, the state of S is determined independently of the state of T and vice versa). We also assume that both signal components are always present in a signal, resulting in four possible component

Strategy	Combination of Components			
	S+T+	S+T-	S-T+	S-T-
Accept All	1	1	1	1
Reject All	0	0	0	0
Follow One Component (S Only)	1	1	0	0
Accept Only S+T+	1	0	0	0
Reject Only S-T-	1	1	1	0

Table 4.1: Accept rates for each of the four component combinations for all five signal following strategies. The final two strategies (Accept Only S+T+, Reject Only S-T-) are the “Confirm the Rare State” strategies. For the Follow One Component strategy, the example of following S only is shown.

combinations: S+T+, S+T-, S-T+, and S-T-. We highlight five signal following strategies, which are defined based on receiver responses to these four combinations. The five strategies are: Accept All, Reject All, Follow One Component, Accept Only S+T+ and Reject Only S-T-. These strategies are illustrated in Table 4.1¹. Two of these strategies (Accept All and Reject All) do not require any signal – subjects perform the same behavior regardless of the states of the signal components. In the Follow One Component strategy, subjects choose one signal component and follow only that one (Table 4.1 shows this strategy for a receiver following only component S). The final two strategies, in contrast, are complex signal following strategies – receivers respond based on the combination of components.

In Rubi and Stephens (*in revision*), we explore how the signaler population influences the optimal receiver strategy. We define the variable p as the proportion of good signalers in the environment (signalers whose true state is good). We show that when the reliabilities q and r are held constant, a receiver’s signal following strategy is determined completely by the value of p . For the purposes of this paper, we set $q = r = 0.8$ (Rubi and Stephens (*in revision*) explores different values of q and r). Figure 4.1 illustrates the strategy payoffs across the full range of p . Unsurprisingly, the Accept All and Reject All strategies are most advantageous when the proportion of good signalers is very high ($p \approx 1$) or very low ($p \approx 0$), respectively. The two complex signal following strategies are most advantageous when one true state is rare, that is, at moderate levels of p . To

¹ For simplicity, we restrict our focus to accept rates of 0 and 1; we expect these idealized strategies to yield maximal benefits.

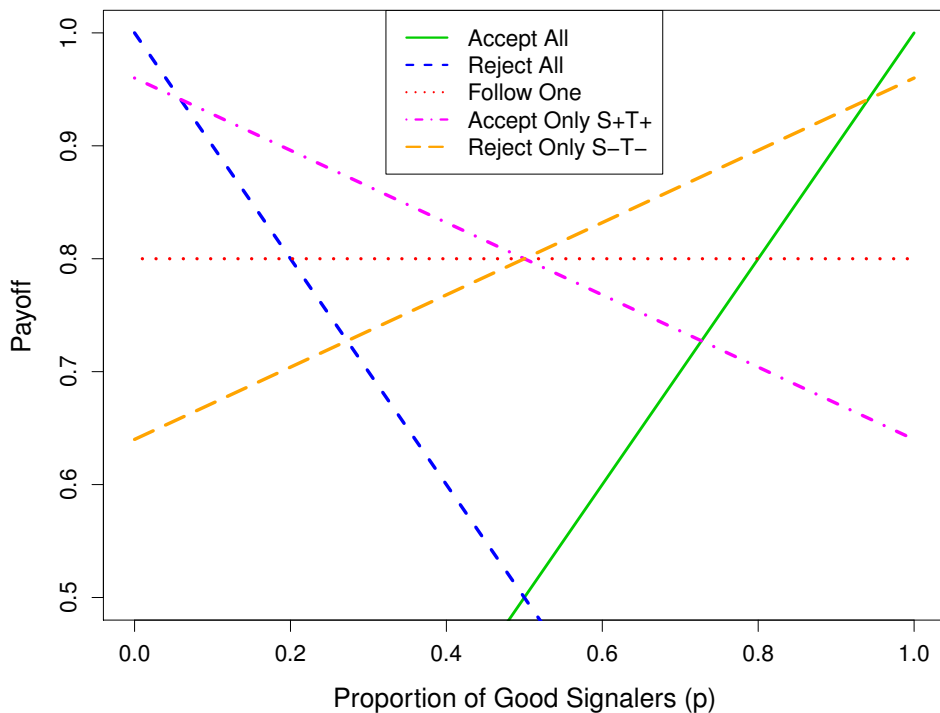


Figure 4.1: Payoffs for the five strategies across a range of p .

summarize these two strategies, receivers generally perform the behavior that matches the more common true state, but switch behavior when both components simultaneously signal that the rare state is true. Accordingly, we refer to these two strategies as Confirm the Rare State strategies. When good and bad signalers are equally common ($p = 0.5$), the payoff of the Follow One Component strategy is equal to the payoffs of both Confirm the Rare State strategies.

In this study, we tested receivers in a lab-based signaling game in which one true state was rare, and assessed responses to determine whether receivers adopted the predicted strategies. Captive blue jays (*Cyanocitta cristata*) served as our experimental subjects. The experimental signals were visual images presented on an LCD screen. Each signal was composed of a square (component S) and a cross (component T). Both shapes were colored, and color was the informative component state. Subjects could accept or

reject the signals by pecking buttons, and were rewarded for matching their action to the true state. We characterized the signal following strategies adopted by receivers at two levels of p , one in which the bad state was rare ($p = 0.8$) and one in which the good state was rare ($p = 0.2$). We first established stable signal following behavior in a pre-treatment at $p = 0.5$ (an equal mix of the good and bad states). Once subjects showed stable signal following behavior (defined in the Treatment Overview section of the Methods), the value of p changed to either 0.2 or 0.8 (depending on the treatment) and the receivers were allowed to act freely. We then examined responses to the four component combinations to characterize the signal following strategies adopted by the subjects.

4.2 Methods

4.2.1 Subjects

Seven adult blue jays of unknown sex, randomly selected from a captive colony, served as the experimental subjects. During the experiment subjects were housed individually in the testing chambers for 23 hours / day (they were removed for an hour for cleaning and maintenance). Subjects ran trials from 7 am to 3 pm daily for the duration of the experiment. They received water ad libitum and were maintained on a 12 hour light:dark cycle. The experiment operated in a closed economy; that is, subjects received all of their food from the experiment, except in cases where a bird's body weight dropped dangerously low (below 85% of its measured ad libitum weight). All housing and experimental procedures were approved by the University of Minnesota Institutional Animal Care and Use Committee (protocol #1408-31752A).

4.2.2 Testing Apparatus

Figure 4.2a depicts an overhead view of the testing apparatus. A perch lever was located at the rear of the box directly below an indicator light. At the front of the box (Figure 4.2b), signals were presented on an LCD screen (Accelelevision LCDP7W). The accept and reject keys (MED Associates ENV-123AM) were located on either side of the screen.

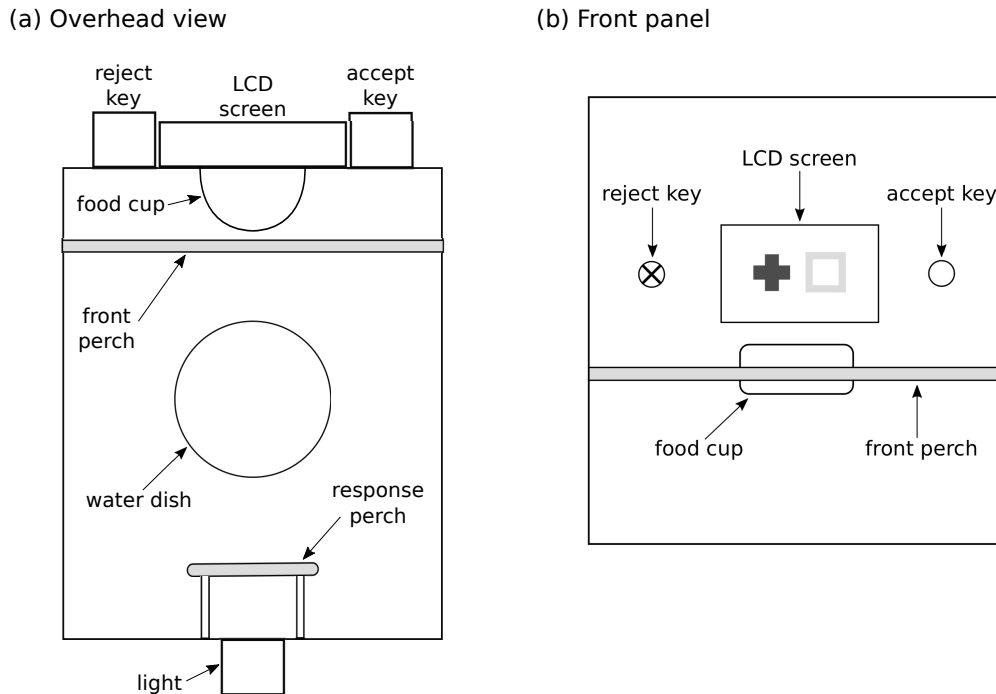


Figure 4.2: An example configuration of the testing apparatus (left-right assignments for the accept and reject keys were randomized across subjects). (a) Overhead view. (b) Front panel.

The reject key was marked with an X and accept keys were unmarked, and the left-right assignments of the accept and reject keys were randomized across subjects. The food cup and magazine light were located directly under the screen. A single stationary perch spanning the width of the box allowed subjects to access the two keys and the food cup. A Med-PC computer program (Med Associates, Burlington, VT, USA) controlled inputs and outputs and recorded data.

4.2.3 Signals

We created two sets of signals using Adobe Illustrator (Adobe Systems, San Jose, CA, USA). Signals consisted of two components, a square and a cross. Color was the informative signal state, and the two shapes were designed such that they covered an equal area on the screen. The two signal sets are shown in Figure 4.3. In signal set 1, the square was either yellow or aquamarine, while the cross was either purple or red.



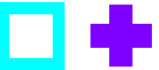





	Combination of Components			
	SaTa	SaTb	SbTa	SbTb
Signal Set 1				
Signal Set 2				

Figure 4.3: The two signal sets. The positive / negative assignments were randomized such that each combination was the S+T+ for at least one subject.

In signal set 2, the square was either orange or green, while the cross was either pink or blue. Treatment order and signal set / treatment pairings were pseudo-randomly assigned across subjects such that every combination of treatment order, treatment, and signal set was tested in at least one subject. Hereafter we refer to the two square signal states as S+ or S- and the two cross signal states as T+ or T-. Positive / negative assignments were randomized such that each combination (e.g., yellow square / red cross) was the S+T+ for at least one subject.

4.2.4 Trial Walkthrough

Trials were separated by an intertrial interval (ITI) of 110 seconds. At the start of each trial, the computer determined the trial type (free or forced) and the state (good or bad) based on the treatment parameters. The computer then independently assigned the signal states (e.g., S+ or S-) probabilistically based on the component reliabilities. To begin the trial, a flashing light at the rear of the box indicated that the trial was ready to begin. The subject began the trial by hopping to the perch lever at the rear of the box. Once the bird occupied the rear perch, the flashing light extinguished and the signal and accept / reject lights illuminated at the front of the box. Subjects responded by pecking either the accept or reject key. Once the subject responded, the computer extinguished the signal and keys and delivered the programmed food reward,

accompanied by a flashing magazine light. The apparatus rewarded correct responses with 2 food pellets (approximately 0.04 g) and delivered no food for incorrect responses. If a subject did not respond in 15 minutes, the trial aborted, the ITI restarted, and the trial was repeated.

4.2.5 Treatment Overview

We arranged trials in blocks of 24, beginning with 4 forced trials followed by 20 free trials. Forced trials (also known as “no choice” trials) are similar to free trials, except that they require subjects to respond in a predetermined way and experience the associated reward; this ensures that subjects experience all possible choices and outcomes in a treatment. For example, in a forced correct accept trial, the computer assigned the component states and the apparatus illuminated the signal and accept / reject keys, exactly as in free trials. However, the trial only ended when the subject pecked the accept key (the reject key was not responsive). The subject received a reward in forced correct trials, but not forced incorrect trials. Trials were grouped into blocks of 24, and the first 4 trials of every block were forced trials: a correct accept, an incorrect accept, a correct reject, and an incorrect reject, presented in a random order. The remaining 20 trials in a block were free trials, in which subjects responded freely and were rewarded according to responses.

Before each treatment, subjects completed a pre-treatment in which the good and bad states were equally common – that is, at $p = 0.5$. The pre-treatment ended when subjects showed stable signal following behavior, which we determined based on two criteria: 1) the overall accept rate for all free trials was greater than 15% and less than 85% (indicating that subjects were following the signal, not simply accepting or rejecting everything), and 2) for three consecutive days, the average free trial accept rate to each of the four component combinations varied by less than 15% (indicating that the accept rates were stable). Once stability was reached in the pre-treatment, the treatment began the following morning. Subjects either began the bad state – rare treatment ($p = 0.8$) or the good state – rare treatment ($p = 0.2$). Between the pre-treatment and the treatment, the only experimental parameter that changed was the p value; all other parameters were kept constant. The treatment ended after subjects completed 1500 free trials; a typical treatment lasted 6-7 days.

4.2.6 Data Analysis

For the analysis, we examined the last 50 free trial responses to each of the four component combinations (S+T+, S+T-, S-T+, S-T-). Recall that we outlined two Confirm the Rare State strategies in the introduction: the Accept Only S+T+ strategy and the Reject Only S-T- strategy (see Table 4.1). We quantified the degree to which our subjects adopted these strategies using an interaction term based on the measured accept rates. We called this term the Strategy Score, and calculated it as follows: $[P(\text{Accept S+T+}) - P(\text{Accept S+T-})] - [P(\text{Accept S-T+}) - P(\text{Accept S-T-})]$. Receivers perfectly following the Accept Only S+T+ strategy or Reject Only S-T- strategy would result in scores of 1 and -1; therefore, we predict a score of 1 in the good state - rare treatment and -1 in the bad state - rare treatment. We calculated the Strategy Score for each bird in each treatment and compared scores in two treatments using a paired t-test.

4.3 Results

4.3.1 Stable signal following behavior

The observed accept rates are shown in Table 4.2. In general, we observed the hypothesized trends, though they were notably weaker than expected. Figure 4.4 shows a box plot of the Strategy Scores for the two treatments. Expected values for each treatment are indicated with the symbol \otimes . A paired t-test indicated that the Strategy Scores were marginally significantly different between the two treatments ($t(6) = 2.478$, $p = 0.048$).

Treatment	Combination of Components			
	S+T+	S+T-	S-T+	S-T-
Good state - rare ($p = 0.2$)	0.23 ± 0.09	0.02 ± 0.01	0.01 ± 0.01	0 ± 0
Bad state - rare ($p = 0.8$)	0.94 ± 0.06	0.97 ± 0.03	0.83 ± 0.09	0.73 ± 0.12

Table 4.2: Observed accept rates for the four component combinations.

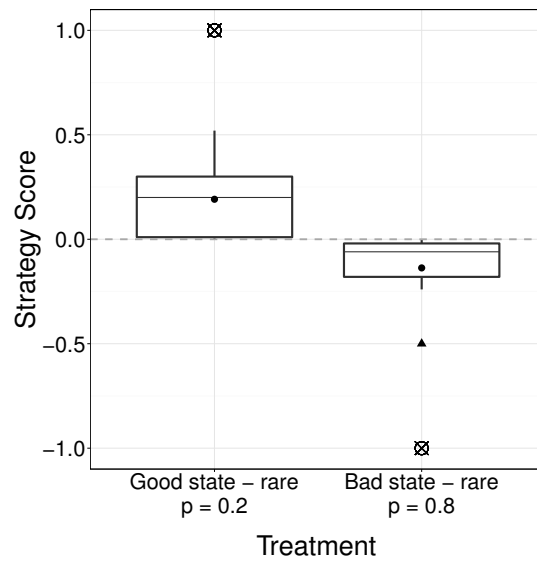


Figure 4.4: Box plot of the strategy Scores in the two treatments. The black dots indicate means, lines indicate medians, box edges span the first to third quartiles, whiskers extend to $1.5 \times$ interquartile range, triangles indicate outliers. The expected values in each treatment are indicated with \otimes .

4.3.2 Transition from Pre-treatment ($p = 0.5$)

Figure 4.5 shows how the accept rates for four component combinations changed over time, and includes the transition from the pre-treatments to the treatments. The data are grouped into blocks of 150. The first 3 data points are the final 3 blocks of the pre-treatment ($p = 0.5$). The next 9 data points show behavior during the first 9 blocks of the treatments. (The total number of trials completed varied slightly across subjects due to differences in response rates, but all subjects completed at least 9 full blocks.) At the start of both treatments, accept rates for all the stimuli move in the same direction (toward zero in the good state – rare treatment and toward one in the bad state – rare treatment). In the good state – rare treatment, the accept rates toward the end of the treatment roughly resemble the hypothesized Accept Only S+T+ strategy, with the crucial difference that the accept rate for S+T+ is much lower than expected. Behavior in the bad state – rare treatment is a bit more muddled, but roughly follows the expected trends as well.

4.4 Discussion

4.4.1 Overview of Results

While we supported our predictions qualitatively, we only found weak quantitative support. The strategy scores in the two treatments differed significantly, and in the expected directions (Figure 4.4). Throughout each treatment, learning curves for the four stimuli roughly followed the expected trajectory, particularly in the good state – rare treatment (Figure 4.5). However, the observed strategies differed from the expected strategies in one critical way; subjects used S+T+ and S–T– combinations to confirm the rare state much less often than expected. Some individuals performed this task better than others, however, even the most extreme individual strategy scores hovered around 0.5 and -0.5, rather than 1 and -1 as expected. Put another way, subjects sometimes did use combinations to confirm the rare state, but had a strong tendency to over-perform the behavior that was more rewarded on average (i.e., they over-rejected in the good state – rare treatment and over-accepted in the bad state – rare treatment).

We could not determine from this study why receivers adopted these weaker versions

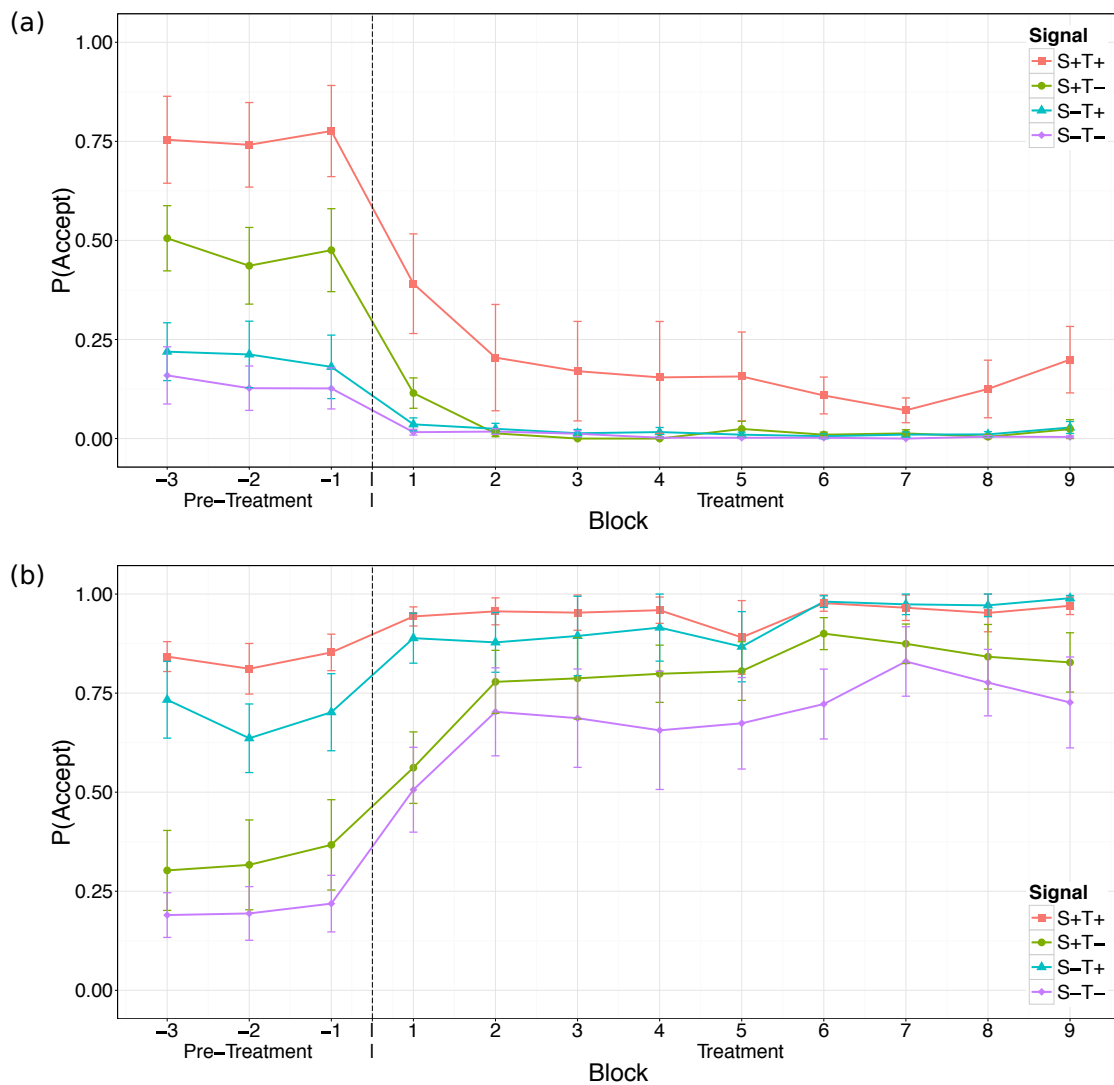


Figure 4.5: Accept rates for the four component combinations. Data are grouped into blocks of 150. These plots show the last 3 blocks of the pre-treatment ($p = 0.5$) and the first nine blocks of the treatments. The dashed lines separate the pre-treatments from the treatments. (a) Good state - rare treatment ($p = 0.2$). (b) Bad state - rare treatment ($p = 0.8$).

of the predicted strategies. One possibility is that receivers had not yet converged on the optimal solutions when the treatment ended. As Figure 4.5 illustrates, it appears that the accept rates might be moving in the expected directions at the end of the treatments (S+T+ is increasing in the good state – rare treatment and S–T– is decreasing in the bad state – rare treatment). Based on previous experiments, we selected what we thought was a conservatively long treatment length; nevertheless, it appears that the observed trends might have been stronger if we had allowed the treatments to run longer. It seems unlikely (though not impossible) that behavior would change dramatically so late in the treatment, however. An alternative interpretation is that this behavior reflects a more general trend we have noted in the jays; they seem to prefer averaging behavior (Accept All / Reject All strategies) to signal following. In this experiment, a subject adopting the Confirm the Rare State strategy would perform the same behavior 80% of the time (accept in the bad state – rare treatment, reject in the good state – rare treatment). Subjects had a tendency to get “stuck” in these behaviors; in other words, they were biased to over-perform the single behavior that was better on average. Several previous studies in our lab have found similar biases; often, the jays adopt these strategies even when they are slightly disadvantageous. Essentially, the jays are willing to pay a cost to not follow a signal. This trend is somewhat intriguing from the perspective of signaling economics; some researchers have argued that assessing signals may be inherently costly for receivers (e.g., Dawkins and Guilford 1991; Jennions and Petrie 1997). Assessment costs such as opportunity costs or increased predation risk are not explicitly applicable to our experimental system, however, they are likely to be general properties of most natural signaling systems. The non-following bias we observe is consistent with the argument that signal following is costly for receivers, however, we can only speculate based on these data.

4.4.2 Rare States and Signaling Systems

Signaling systems in which one signaler state is rare are likely to be common in nature. For example, high-quality males might be rarer than low-quality males (sexual signaling), toxic prey might be rarer than palatable prey (aposematic signaling), and good fighters might be rarer than poor fighters (agonistic signaling). Of course, these examples highlight some of the challenges of this approach; classifying signalers, signals,

and even receiver responses into discrete categories can be challenging. Some signaling systems will be easier to classify than others; in an ideal case, signaler states would be binary and clearly delineated. Identity signals, which facilitate individual recognition (Johnstone 1997; Tibbetts and Dale 2007), are a clear and compelling example. A given signaler is either “individual X” or “not individual X,” making classification straightforward. In addition, there is only one “individual X” in the population, so the true state “individual = X” will always be rare (specifically, $p = 1/n$, where n is the social group size). The need for multiple components in identity signals is intuitive; at a minimum, individuals must vary in enough qualities to be identified as unique (Beecher 1989; Dale et al. 2001; Tibbetts 2004; Tibbetts and Dale 2007). Identity signals are diverse, but multicomponency appears to be ubiquitous (e.g., Dale et al. 2001; DelBarco-Trillo et al. 2012; Proops et al. 2008). Though classifying some types of signaling systems might be more challenging, we suggest that this model is worth exploring in any system with asymmetries between signaler states.

4.4.3 Concluding Summary

Overall, these data broadly support our prediction that receivers will use multiple signal components in environments in which one signaler state is less common. Our subjects roughly adopted the predicted “Confirm the Rare State” behavioral strategies, though the trends were somewhat weaker than expected. Our results suggest that our model has merit in predicting complex signal following behavior, however, more work needs to be done to confirm these trends.

Chapter 5

Detection of signaling *Schizocosa ocreata* by an avian predator¹

5.1 Introduction

Detectability is a critical determinant of the effectiveness of a signal, and as such, signals are under selection to maximize the probability of detection against noisy backgrounds (Guilford and Dawkins 1991, 1993; Wiley 2006). One of the most commonly-cited benefits of complex signals is that they are more detectable in environmental noise (Candolin 2003; Hebets and Papaj 2005; Partan and Marler 2005; Rowe 1999). A caveat of highly detectable signals, however, is that they are likely also more detectable to predators. It is well-established that predators exploit the signals of their prey, and that signalers face a trade-off between the benefits of successful communication and the costs of increased conspicuousness to predators (Magnhagen 1991; Zuk and Kolluru 1998). For signalers with multiple predator species, the costs of complex signaling may be even more pronounced. Complex signals are likely to be perceptible by a wider range of predators simply because there are more components to exploit. This is especially true of multimodal signals, which will inevitably be conspicuous to more predator species because they occur across multiple sensory modalities (indeed, this is a proposed benefit of multimodal aposematic signals – Ratcliffe and Nydam 2008; Rowe and Halpin

¹ This experiment was a collaborative project with Drs. George Uetz (University of Cincinnati) and Dave Clark (Alma College).

2013). Despite the logic of this argument, the literature still lacks an explicit discussion of complex signals and eavesdropping by multiple predator species. The experiment presented in this chapter contributes to a broad and comprehensive research program aimed at tackling this topic in a focal organism, the brush-legged wolf spider (*Schizocosa ocreata*).

S. ocreata males perform an elaborate multimodal courtship display. Males signal visually by tapping, waving, and arching forelegs ornamented with dark tufts, as well as seismically by drumming their legs and abdomens on the substrate and stridulating using an organ located in the pedipalps (Stratton and Uetz 1983, 1986; Uetz and Denterlein 1979). Female *S. ocreata* prefer males with large tufts and high amplitude seismic displays (Gibson and Uetz 2008; McClintock and Uetz 1996; Uetz and Norton 2007), both of which are honest indicators of quality (Gibson and Uetz 2012; Uetz et al. 2002). Visual or seismic components alone are sufficient to elicit responses from females, however, females are quicker to respond and are more receptive when displays are multimodal (Uetz et al. 2009). Predation studies have found substantial overlap in the traits that attract females and those that attract predators. American toads (*Bufo americanus*) detect male *S. ocreata* more quickly when leg tufts are large (Roberts and Uetz 2008). In *Hogna helluo*, a larger species of wolf spider, leg waving and tuft size interact to increase the detectability of signaling males (Pruden and Uetz 2004). The jumping spider *Phidippus clarus* is sensitive to leg motion, tuft size, and seismic components, responding most strongly when all are present (Roberts et al. 2006). Our aim in the experiment was to test detection of signaling *S. ocreata* by an avian predator. Birds differ markedly in sensory abilities and predation behavior from the previous predators tested. In addition, previous experiments focused on predators that view *S. ocreata* from a lateral perspective – this experiment is the first to examine detection of signaling *S. ocreata* from an overhead perspective.

We also included a body color manipulation in this experiment. Though *S. ocreata* show little sexual dichromatism (suggesting that coloration is not a component of the sexual signal), coloration is likely to play a key role in detection by females and predators. *S. ocreata* have a mostly dark body with a lighter stripe down the dorsum. Spectral reflectance analyses indicate that *S. ocreata* as viewed from a lateral perspective show high contrast with leaf litter backgrounds (Clark et al. 2011). This dark coloration

improves detection by females against natural backgrounds (Uetz et al. 2010). The spectral reflectance of dorsal coloration, in contrast, overlaps significantly with that of a leaf litter background (Clark et al. 2011). This suggests that dorsal coloration is cryptic from an overhead perspective. The aim of this manipulation was to compare the detectability of natural dorsal coloration with that of an idealized background-matching coloration. We generated the latter stimuli by altering spiders to match the average spectral reflectance of the leaf litter background.

Our experimental stimuli were video playbacks of digitized *S. ocreata* recorded from an overhead orientation. We tested three levels of behavior (courting, walking, or stationary) and two levels of body coloration (natural or background-matching). The behavior manipulation tested whether the visual aspects of courtship behavior influenced detection rates. We focused only on the visual behavioral aspect of the courtship signal because avian predators are unlikely to perceive seismic components (due to their distance from the substrate) or leg tufts (due to the size and orientation of the tufts). The body color manipulation tested whether natural coloration functions as a background-matching cryptic strategy.

We used blue jays (*Cyanocitta cristata*) as our experimental subjects. Blue jays were a fitting choice for this experiment; blue jays co-occur with *S. ocreata* and are generalist predators of arthropods, and *S. ocreata* show antipredatory behaviors (freezing and increased latency to resume courtship) when presented with playbacks of blue jay vocalizations (Lohrey et al. 2009). We presented subjects with video playbacks in a “present / absent” detection task. We first trained subjects to identify *S. ocreata* and to peck buttons to register “present” and “absent” responses. In this pre-experiment training, subjects were trained to identify all six categories of spider stimuli on a neutral gray background. In the experiment itself, subjects were tested on the same six spider categories, but the spiders were superimposed onto a leaf litter background. We presented subjects with video playbacks at regular intervals and they pecked buttons to indicate whether a spider was present or absent on the screen. They were rewarded with food for correct responses. All categories of spider-present and spider-absent stimuli were pseudo-randomly shuffled together, and we tracked the detectabilities and response latencies for each type of stimulus over the course of the experiment.

5.2 Methods

5.2.1 Subjects

Nine adult blue jays of unknown sex, randomly selected from a captive colony, served as our experimental predators. During the experiment subjects were housed individually in the testing chambers for 23 hours / day (they were removed for an hour for cleaning and maintenance). Subjects ran trials from 7 am to 3 pm daily for the duration of the experiment. They received water ad libitum and were maintained on a 12 hour light:dark cycle. The experiment operated in a closed economy; that is, subjects received all of their food from the experiment. To ensure our subjects' well being, we provided additional food in cases where a bird's body weight dropped below 85% of its measured ad libitum weight. All housing and experimental procedures were approved by the University of Minnesota Institutional Animal Care and Use Committee (protocol #1408-31752A).

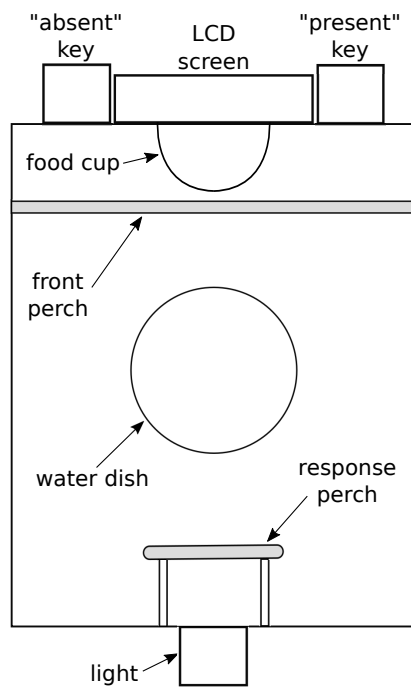
5.2.2 Testing Apparatus

Figure 5.1a depicts an overhead view of the testing apparatus. A perch lever was located at the rear of the box directly below an indicator light. At the front of the box (Figure 5.1b), stimuli were presented on an LCD screen (Accelelevision LCDP7W). The "present" and "absent" keys (standard pigeon keys – MED Associates ENV-123AM) were located on either side of the screen. The absent keys were marked with an X and present keys were unmarked, and the left-right position of the two keys was randomized across subjects. The food cup and magazine light were located directly under the screen. A single stationary perch spanning the width of the box allowed subjects to access the two keys and the food cup. A Med-PC computer program (Med Associates, Burlington, VT, USA) controlled inputs and outputs and recorded data.

5.2.3 Video Stimuli

The spider-present stimuli were high contrast video playbacks of brush-legged wolf spiders (*Schizocosa ocreata*) superimposed on an image of a leaf litter background. Spider-absent stimuli were leaf litter backgrounds with no spider present. We tested three levels of behavior: courting, walking, and stationary. In both courting and walking

(a) Overhead view



(b) Front panel

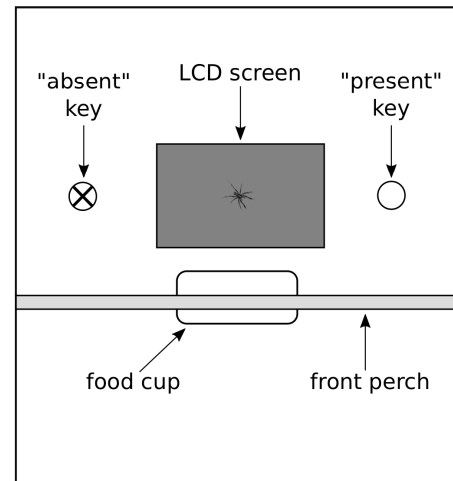


Figure 5.1: An example schematic of the testing apparatus. The positions of the "present" and "absent" keys were randomized across subjects. (a) Overhead view. (b) Front panel.

videos, playbacks began with spiders in different starting locations and orientations, and the spiders moved across the screen. If the spider reached the edge of the screen before the subject responded, it moved offscreen and then immediately reappeared from a different location along the edge of the screen. In all stimuli, the spider was fully on-screen for at least 4 seconds at the beginning of the stimulus presentation. In stationary videos, location and orientation varied but the spiders were motionless. Stationary spiders were always completely onscreen. We also tested two levels of coloration: natural coloration and idealized background-matching coloration. The background-matching spiders were artificially colored to match the average reflectance of the leaf litter background. Natural-colored spiders were not modified. Figure 5.2 shows four stationary spider-present stimuli. The neutral gray background was used in the pre-experiment training and the leaf litter background was used in the experiment. Both natural and background-matching spiders are shown on each background. Spiders are indicated with an arrow. Walking and courting videos looked very similar to the stimuli in Figure 5.2, but the spider moved across the screen while performing the appropriate behavior.

Due to constraints on generating and presenting stimuli, it was necessary to repeat stimuli throughout the experiment. For each stimulus category (e.g., courting / background-matching) we generated four different videos. We randomized the videos such that the starting position of each spider, both within and between stimulus categories, was unique. In addition, we presented versions of each video that had been rotated 180 degrees. That is, the spider could be in any of 48 unique positions at the start of the playback (eight different positions for each of the six stimulus categories). The blank leaf litter backgrounds (spider-absent) were generated to match the spider-present stimuli and presented at equal frequency; in other words, for each presentation of a given spider-present stimulus there was a presentation of an identical leaf litter background with no spider present.

5.2.4 Trial Walkthrough

Trials were separated by an intertrial interval (ITI) of 110 seconds. When the ITI had elapsed, the computer determined the trial type (spider-present or spider-absent) and the stimulus. A flashing light at the rear of the box then indicated that the trial was ready to begin. Subjects began the trial by hopping to a perch lever at the rear of the

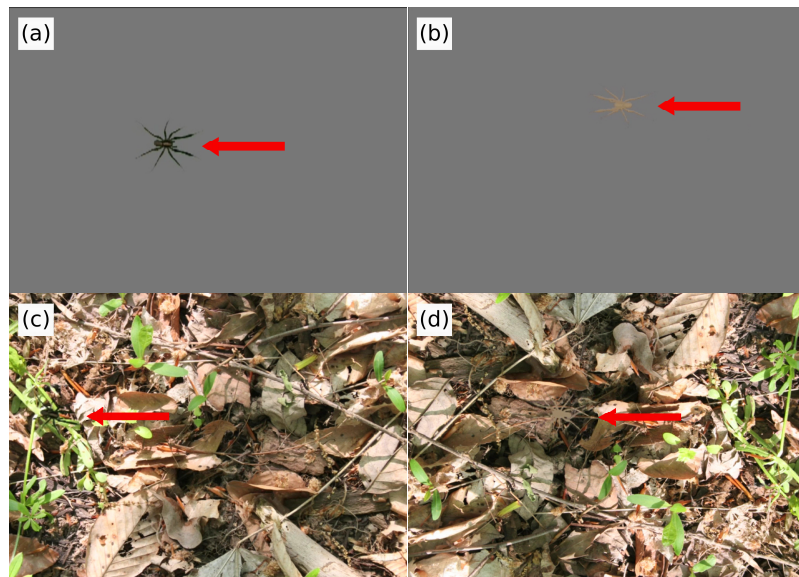


Figure 5.2: Example stationary spider stimuli. Courting and walking spiders looked very similar, but the spider was performing the appropriate motion. Neutral gray backgrounds were used in the pre-experiment training (panels a and b), and leaf litter backgrounds were used in the experiment (panels c and d). Blank spider-absent stimuli looked identical, but with no spider present. The spider is located to the left of the arrow. (a) A natural-colored spider on a neutral gray background. (b) A background-matching spider on a neutral gray background. (c) A natural-colored spider on a leaf litter background. (d) A background-matching spider on a leaf litter background.

box, causing the flashing light to extinguish, the two response keys to illuminate, and the video playback to begin. Once the video playback started, the timer to measure the reaction time began. Subjects responded by pecking either the present key or the absent key. Once the subject responded, the computer extinguished both response keys, ended the video playback, and delivered a food reward if applicable (accompanied by a flashing magazine light). If a subject did not respond in 9 minutes, the trial aborted, the ITI restarted, and the trial was repeated.

5.2.5 Experiment Progression

Prior to the start of the experiment, subjects were trained to identify all six categories of spider stimuli on a neutral gray background. Subjects moved on to the experiment once they had reach a detection threshold for all stimuli. The detection threshold was defined as when the proportion of correct present responses exceeded the proportion of incorrect present responses by at least 50% (following the terminology outlined in the Data Analysis section: $P(\text{Hit}) - P(\text{False Alarm}) \geq 0.5$). All birds readily achieved this criterion, indicating that they were able to detect all the categories of spider stimuli on a neutral gray background.

Once the pre-experiment training ended, subjects started the experiment the following morning. The experiment continued until 2,500 total trials had been completed, typically over the course of 12 - 13 days. For the analysis, these trials were arranged into blocks of 250, for a total of 10 blocks. The computer presented the stimuli in a psuedo-random order; the randomization was restricted in order to balance the number of presentations of spider-present / spider-absent stimuli, the different stimulus categories, and the different stimulus versions in each category.

5.2.6 Data Analysis

We assessed the detectability of the various spider stimuli using the parameter d' from signal detection theory (Tanner and Swets 1954). In a presence / absence task such as this one, d' is calculated based on the proportion of “hits” and “false alarms.” A hit is a correct present response; that is, a spider is present and the subject correctly responds that a spider is present. A false alarm is an incorrect present response; that

is, no spider is present on the screen but the subject incorrectly responds that a spider is present. These two measures fully describe performance on a presence / absence task. The parameter d' is the difference in the z scores of these two probabilities; that is, it is calculated: $d' = z(\text{hit rate}) - z(\text{false alarm rate})$. We first applied a loglinear adjustment to the hit and false alarm counts to correct for extreme values (following Stanislaw and Todorov 1999, p. 143-4) – this is necessary when hit and false alarm rates of 0 and 1 arise since the z scores are $-\infty$ and ∞ , respectively. Using these adjusted rates, we then calculated d' for each stimulus category in each block and used it to assess how the stimuli changed in detectability over the course of the experiment. To characterize the final detectabilities at the end of the experiment, we examined data from the last three blocks.

We also measured the reaction time to a subset of stimuli. For this measure, we only looked at data from hits (correct present responses), and only to walking and courting spiders; the hit rate for stationary spiders was very low, resulting in too few data points to characterize the reaction time. The clock timing reaction time started when the video playback appeared on the screen and stopped when the subject pecked a button to register a response. We examined reaction time over the last three blocks.

5.3 Results

5.3.1 Detectability

Figure 5.3 shows the detectability (quantified using d') of all spider categories in the last three blocks. The sphericity assumption was violated, so we applied a Greenhouse-Geisser correction. A repeated measures ANOVA with factors of behavior and coloration revealed a significant coloration*behavior interaction ($F_{2,16} = 8.04$, $p = 0.02$). We used Tukey post-hoc comparisons to determine significant differences. Unsurprisingly, stationary spiders were much less detectable than moving spiders. In stationary spiders and courting spiders, natural and background-matching coloration was similarly detectable. However, coloration affected detectability in walking spiders; natural-colored walking spiders were as detectable as courting spiders, but background-matching walking spiders were significantly less detectable. These trends were reflected over the entire course of the experiment, as well. Figure 5.4 tracks the detectability of the stimulus categories

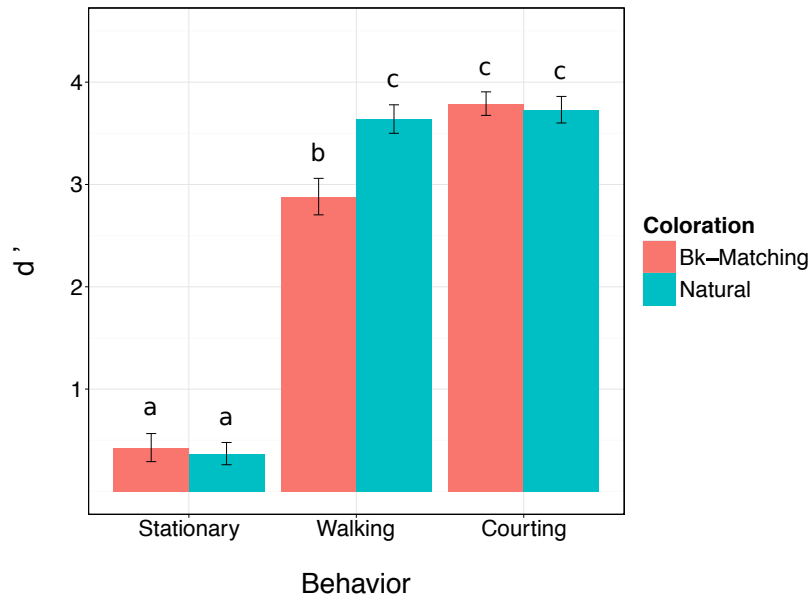


Figure 5.3: Detectability (measured by d') in the last three blocks. Bk-Matching stands for background-matching.

over all blocks. A repeated measures ANOVA with factors of behavior, coloration, and block revealed significant significant interactions of behavior*coloration ($F_{2,16} = 15.71$, $p < 0.001$), as well as behavior*block ($F_{2,16} = 9.82$, $p = 0.002$).

5.3.2 Reaction Time

Figure 5.5 shows the reaction time to courting and walking spiders in the last three blocks. A repeated measures ANOVA revealed that courting spiders were detected significantly faster than walking spiders ($F_{1,8} = 10.22$, $p = 0.013$). The coloration*behavior interaction and the main effect of coloration were not significant.

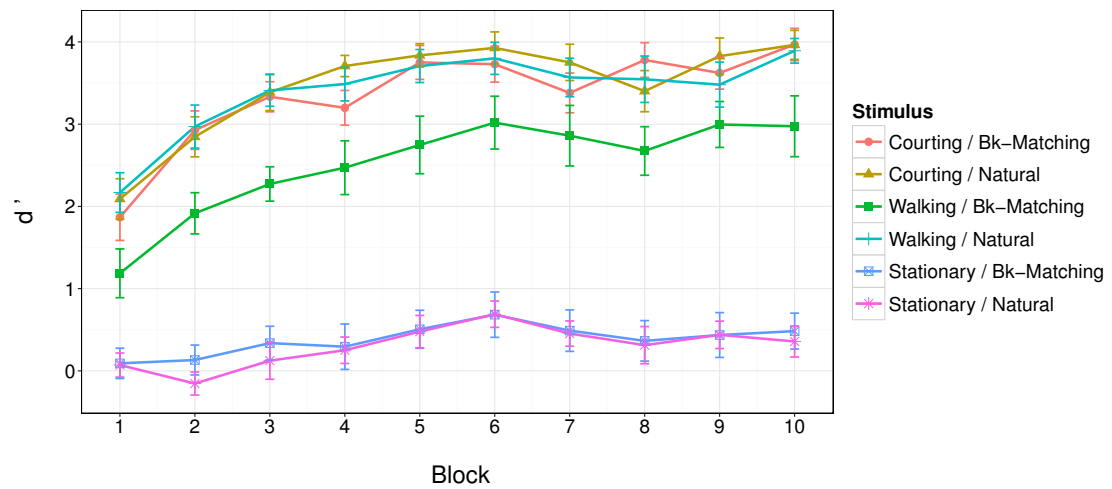


Figure 5.4: Detectability (measured by d') of all stimulus types over the course of the entire experiment. Bk-Matching stands for background-matching coloration.

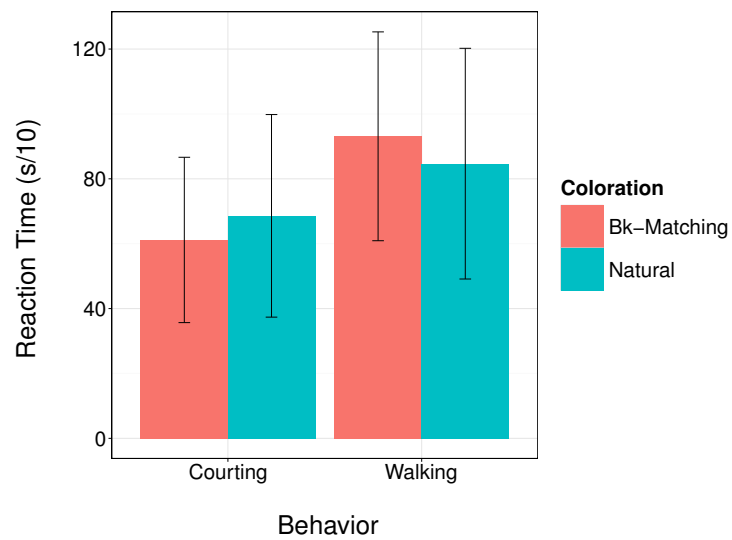


Figure 5.5: Reaction time for courting and walking spiders in the last three blocks when the response was a hit. Bk-Matching stands for background-matching.

5.4 Discussion

5.4.1 Results in Context

Perhaps the most striking effect we found was that our subjects were highly attuned to motion; stationary spiders were much less detectable than either walking or courting spiders. Birds have excellent visual acuity so motion may be a critical prey cue, particularly when searching for prey against a visually complex background. While the observation that stationary targets are less detectable than moving targets may seem mundane, it is ecologically relevant to this system. Freezing is an important anti-predation strategy in *S. ocreata*. Spiders freeze in response to acute predator cues and decrease locomotion in perceived high-predation environments (Fowler-Finn and Hebets 2011; Lohrey et al. 2009). Lohrey et al. (2009) focused on avian predator cues specifically and found that both simulated beak-pecking and acoustic playbacks of bird calls stimulated freezing behavior. These authors also compared responses of *S. ocreata* to playbacks of blue jays (which are predatory) and mourning doves (which are non-predatory), and found that blue jay calls elicited a greater freezing response. Our results suggest that freezing is probably a highly effective anti-predation strategy for *S. ocreata* when faced with avian predators.

We found some support for the hypothesis that courting spiders are more vulnerable to predation than walking spiders. The relative detectability of courting and walking spiders was influenced by coloration (see the discussion below), however, courting spiders in general were detected significantly more quickly than walking spiders. There are various reasons that courting spiders might be more visually conspicuous. For example, courting spiders adopt a more winding path, and move their abdomens as well as their legs. However, given the importance of motion discussed above, a particularly compelling explanation is that courting spiders spend more time in motion than walking spiders. The locomotion of walking spiders can be described as forward movement along a path, punctuated by periodic stopping. Courting spiders follow a similar pattern of locomotion, however, the forelegs of a courting spider are almost constantly in motion. When a walking spider is stopped, it looks very similar to a frozen spider; in contrast, a courting spider is almost never completely motionless. It is interesting to note that

while courting spiders were detected significantly faster, they weren't necessarily detected more often; walking natural-colored spiders had similar d' values to courting spiders. Detection speed and detectability are likely to be less independent in nature. For example, a walking spider may have more time to detect the presence of a predator and freeze, resulting in reduced detection rates. Nevertheless, this result highlights the fact that avian predators are attuned to motion, making locomotion of any kind highly detectable.

The coloration of the spiders had differing effects based on behavior. When spiders were stationary or courting, detection rates did not vary based on coloration. In contrast, when spiders were walking, natural-colored spiders were significantly more detectable than background-matching spiders. These data suggest that the natural coloration of *S. ocreata* often functions similarly to a background-matching strategy, however, this effect is contingent upon behavior. We can only speculate on the specific mechanism of this effect based on these data. For example, it may be that stationary natural-colored spiders are actually more detectable than stationary background-matching spiders, but only when motion cues draw the subject's focus to the appropriate area on the screen. As described above, "walking" spiders spend some time being stationary; it is conceivable that the jays catch a glimpse of motion before a walking spider stops, then are able to locate the stationary spider once they know where to look. In contrast, the jays might not be likely to find purely stationary spiders since they must scan the entire screen.

Whatever the mechanism, the results suggest that natural coloration does not function quite as effectively as true background-matching coloration. The spectral reflectance of natural dorsal coloration overlaps significantly with that of a leaf litter background (Clark et al. 2011). This is probably due to the broad, light-colored dorsal stripe; however, the dark legs and sides are also visible from an overhead perspective. As discussed in the introduction, the lateral coloration of *S. ocreata* shows high contrast with leaf litter backgrounds (Clark et al. 2011), and this contrast enhances detectability by females (Uetz et al. 2010). It is particularly critical that courting males are correctly identified by females because they are subject to pre-copulatory sexual cannibalism (Persons and Uetz 2005). Dorsal coloration in *S. ocreata* may be the "best-case scenario" for cryptic coloration, as it affords *S. ocreata* some protection from overhead

predators while still being conspicuous to females from a lateral perspective.

5.4.2 Multi-Species Predation and Complex Signals

Little work to date focuses on multi-species predation and complex signaling. More attention has been paid to detection of complex signals by individual predators, and there is good evidence that they can benefit from complex signals in much the same way that receivers do – namely, improved detection and localization (e.g., Bernal et al. 2006; Halfwerk et al. 2014a; Page and Ryan 2008; Rhebergen et al. 2015; Roberts et al. 2006). A small amount of work has been done on multi-species predation and aposematic signals, a topic closely related to multi-species eavesdropping but with very different implications for selection on signals (see Pearson 1989; Rowe and Halpin 2013 for general discussions). Heberts and Papaj (2005) suggested that perceptual variability in receivers could select for complex signals, though they focused on intraspecific variability; as Rowe and Halpin (2013) point out, this hypothesis is much more compelling for aposematic signals that may target multiple receiver species. Ratcliffe and Nydam (2008) used a phylogenetic approach to demonstrate that multi-species predation pressure is likely to select for the evolution of multimodal aposematic signals, however, there is little direct evidence that such signals target multiple predator species. An exception is found in the Peacock Butterfly (*Inachis io*). This species emits a wing-flick defensive display; the wing motion reveals hidden eyespots and produces both a hissing and a clicking sound (Olofsson et al. 2012). Vallin et al. (2005) showed that avian predators avoided visual, but not acoustic, components of the display. Subsequently, Olofsson et al. (2012) found that nocturnal mouse predators tested in complete darkness were deterred by the display, suggesting they reacted to the acoustic component. Aposematic species with both diurnal and nocturnal predators may be a good starting point for future research; some authors have suggested that the combination of diurnal and nocturnal predators could select for multimodal signals (Ratcliffe and Nydam 2008; Rowe and Halpin 2013), however, more research needs to be done in this area.

The literature on complex signaling and eavesdropping by multiple predator species is even more sparse; in both cases, the logistic difficulty of testing several predator species might be to blame. Comprehensive research programs focusing on a single species may provide the best data to date. For example, John Endler's expansive work

on Trinidad guppies (*Poecilia reticulata*) has demonstrated that predator communities affect a variety of male traits, including the size and composition of color spots that are used in both courtship and crypsis (e.g., Endler 1991, 1995; Reznick and Endler 1982). Mike Ryan and colleagues have conducted extensive research on sexual signaling in túngara frogs, and have described eavesdropping by both bats and parasitic flies. Flies use components of the courtship call to localize frogs (Bernal et al. 2006), while bats use a combination of calls, vocal sac movement, and even water ripples produced in association with calling (Halfwerk et al. 2014b; Page and Ryan 2008; Rhebergen et al. 2015). The work presented in this chapter contributes to a comprehensive research program by George Uetz and colleagues that has shown that the various components of the *S. ocreata* sexual signal are subject to eavesdropping by a diverse array of predators. Toads and wolf spider predators exploit leg ornaments, birds and wolf spider predators exploit the leg waving display, and jumping spider predators exploit all those traits as well as seismic signal components (Pruden and Uetz 2004; Roberts et al. 2006; Roberts and Uetz 2008). Such cross-experimental comparisons provide valuable insights, however, more work needs to be done. A first step will be identifying focal signaler species; ideal species will have elaborate (probably multimodal) displays and will be subject to predation by diverse taxa.

5.4.3 Concluding Summary

Our experimental eavesdroppers were highly sensitive to motion; walking and courting spiders were much more detectable than stationary spiders. Subjects detected courting and walking spiders at similar rates, however, they detected courting spiders more quickly. Though spiders with natural coloration were usually equally detectable to spiders with an idealized background-matching coloration, natural-colored spiders were more detectable when they were walking. The reason for this interaction between coloration and behavior is unknown, however, it suggests that natural coloration is not quite as cryptic as true background-matching coloration would be.

References

- Akre, K. L. and Ryan, M. J. (2010). Complexity increases working memory for mating signals. *Current Biology*, 20(6):502–5.
- Alonso-Alvarez, C., Bertrand, S., Devevey, G., Gaillard, M., Prost, J., Faivre, B., and Sorci, G. (2004). An experimental test of the dose-dependent effect of carotenoids and immune activation on sexual signals and antioxidant activity. *The American Naturalist*, 164(5):651–659.
- Alvarado, J. C., Vaughan, J. W., Stanford, T. R., and Stein, B. E. (2007). Multisensory versus unisensory integration: contrasting modes in the superior colliculus. *Journal of Neurophysiology*, 97:3193–3205.
- Amorim, M. C. P., Pedroso, S. S., Bolgan, M., Jordão, J. M., Caiano, M., and Fonseca, P. J. (2013). Painted gobies sing their quality out loud: acoustic rather than visual signals advertise male quality and contribute to mating success. *Functional Ecology*, 27(2):289–298.
- Andersson, S., Pryke, S. R., Ornborg, J., Lawes, M. J., and Andersson, M. (2002). Multiple receivers, multiple ornaments, and a trade-off between agonistic and epigamic signaling in a widowbird. *The American Naturalist*, 160(5):683–91.
- Aronsson, M. and Gamberale-Stille, G. (2008). Domestic chicks primarily attend to colour, not pattern, when learning an aposematic coloration. *Animal Behaviour*, 75(2):417–423.
- Aronsson, M. and Gamberale-Stille, G. (2012). Colour and pattern similarity in mimicry:

- evidence for a hierarchical discriminative learning of different components. *Animal Behaviour*, 84(4):881–887.
- Ballentine, B., Searcy, W. A., and Nowicki, S. (2008). Reliable aggressive signalling in swamp sparrows. *Animal Behaviour*, 75(2):693–703.
- Beecher, M. D. (1989). Signalling systems for individual recognition: an information theory approach. *Animal Behaviour*, 38(2):248–261.
- Bernal, X. E., Rand, A. S., and Ryan, M. J. (2006). Acoustic preferences and localization performance of blood-sucking flies (*Corethrella* Coquillett) to túngara frog calls. *Behavioral Ecology*, 17(5):709–715.
- Blount, J. D., Speed, M. P., Ruxton, G. D., and Stephens, P. A. (2009). Warning displays may function as honest signals of toxicity. *Proceedings of the Royal Society B: Biological Sciences*, 276(1658):871–877.
- Blumstein, D. T., Verneyre, L., and Daniel, J. C. (2004). Reliability and the adaptive utility of discrimination among alarm callers. *Proceedings of the Royal Society B: Biological Sciences*, 271(1550):1851–1857.
- Bradbury, J. and Vehrencamp, S. (2000). Economic models of animal communication. *Animal Behaviour*, 59(2):259–268.
- Bro-Jørgensen, J. (2010). Dynamics of multiple signalling systems: animal communication in a world in flux. *Trends in Ecology and Evolution*, 25(5):292–300.
- Bro-Jørgensen, J. and Dabelsteen, T. (2008). Knee-clicks and visual traits indicate fighting ability in eland antelopes: multiple messages and back-up signals. *BMC Biology*, 6(1):47.
- Brower, L. P. (1969). Ecological chemistry. *Scientific American*, 220(2):22–9.
- Brown, S. G., Boettner, G. H., and Yack, J. E. (2007). Clicking caterpillars: acoustic aposematism in *Antheraea polyphemus* and other Bombycoidea. *The Journal of Experimental Biology*, 210(6):993–1005.

- Bura, V. L., Fleming, A. J., and Yack, J. E. (2009). What's the buzz? Ultrasonic and sonic warning signals in caterpillars of the great peacock moth (*Saturnia pyri*). *Die Naturwissenschaften*, 96(6):713–8.
- Candolin, U. (2003). The use of multiple cues in mate choice. *Biological Reviews of the Cambridge Philosophical Society*, 78(4):575–95.
- Claridge, M. (1974). Stridulation and defensive behaviour in the ground beetle, *Cychrus caraboides*. *Journal of Entomology Series A, General*, 49(1):7–15.
- Clark, D. L., Roberts, J. A., Rector, M., and Uetz, G. W. (2011). Spectral reflectance and communication in the wolf spider, *Schizocosa ocreata* (Hentz): simultaneous crypsis and background contrast in visual signals. *Behavioral Ecology and Sociobiology*, 65(6):1237–1247.
- Clarke, J. C., Westbrook, R. F., and Irwin, J. (1979). Potentiation instead of overshadowing in the pigeon. *Behavioral and Neural Biology*, 25(1):18–29.
- Cohen, S., Stebbins, W., and Moody, D. (1978). Audibility thresholds of the Blue Jay. *The Auk*, 95(1):563–568.
- Cook, S. and Roper, T. (1989). Responses of chicks to brightly coloured insect prey. *Behaviour*, 110(1):276–293.
- Costanzo, K. and Monteiro, A. (2007). The use of chemical and visual cues in female choice in the butterfly *Bicyclus anynana*. *Proceedings of the Royal Society B: Biological Sciences*, 274(1611):845–51.
- Couvillon, P. a. and Bitterman, M. E. (1982). Compound conditioning in honeybees. *Journal of Comparative and Physiological Psychology*, 96(2):192–199.
- Dale, J., Lank, D., and Reeve, H. (2001). Signaling individual identity versus quality: a model and case studies with ruffs, queleas, and house finches. *The American Naturalist*, 158(1):75–86.
- Dawkins, M. S. and Guilford, T. (1991). The corruption of honest signalling. *Animal Behaviour*, 41(5):865–873.

- Deag, J. and Scott, G. (1999). “Conventional” signals in avian agonistic displays: integrating theory, data and different levels of analysis. *Journal of Theoretical Biology*, 196(2):155–162.
- DelBarco-Trillo, J., Sacha, C. R., Dubay, G. R., and Drea, C. M. (2012). Eulemur, me lemur: the evolution of scent-signal complexity in a primate clade. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1597):1909–22.
- Dunlap, A. S. and Stephens, D. W. (2014). Experimental evolution of prepared learning. *Proceedings of the National Academy of Sciences*, 111(32):11750–11755.
- Elias, D., Hebets, E., Hoy, R., and Mason, A. (2005). Seismic signals are crucial for male mating success in a visual specialist jumping spider (Araneae: Salticidae). *Animal Behaviour*, 69(4):931–938.
- Endler, J. A. (1991). Variation in the appearance of guppy color patterns to guppies and their predators under different visual conditions. *Vision Research*, 31(3):587–608.
- Endler, J. A. (1995). Multiple-trait coevolution and environmental gradients in guppies. *Trends in Ecology and Evolution*, 10(1):22–29.
- Endler, J. A. and Basolo, A. (1998). Sensory ecology, receiver biases and sexual selection. *Trends in Ecology & Evolution*, 13(10):415–420.
- Fowler-Finn, K. D. and Hebets, E. A. (2011). The degree of response to increased predation risk corresponds to male secondary sexual traits. *Behavioral Ecology*, 22(2):268–275.
- Gamberale-Stille, G. and Guilford, T. (2004). Automimicry destabilizes aposematism: predator sample-and-reject behaviour may provide a solution. *Proceedings of the Royal Society B: Biological Sciences*, 271(1557):2621–2625.
- Gibson, J. S. and Uetz, G. W. (2008). Seismic communication and mate choice in wolf spiders: components of male seismic signals and mating success. *Animal Behaviour*, 75(4):1253–1262.

- Gibson, J. S. and Uetz, G. W. (2012). Effect of rearing environment and food availability on seismic signalling in male wolf spiders (Araneae: Lycosidae). *Animal Behaviour*, 84(1):85–92.
- Gingras, G., Rowland, B. A., and Stein, B. E. (2009). The differing impact of multisensory and unisensory integration on behavior. *The Journal of Neuroscience*, 29(15):4897–4902.
- Grafe, T. U., Preininger, D., Sztatecsny, M., Kasah, R., Dehling, J. M., Proksch, S., and Hödl, W. (2012). Multimodal communication in a noisy environment: a case study of the Bornean rock frog *Staurois parvus*. *PloS one*, 7(5):e37965.
- Grether, G. F., Kolluru, G. R., and Nersissian, K. (2004). Individual colour patches as multicomponent signals. *Biological Reviews of the Cambridge Philosophical Society*, 79(3):583–610.
- Guilford, T. and Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour*, 42(1):1–14.
- Guilford, T. and Dawkins, M. S. (1993). Receiver psychology and the design of animal signals. *Trends in Neurosciences*, 16(11):430–436.
- Halfwerk, W., Dixon, M. M., Ottens, K. J., Taylor, R. C., Ryan, M. J., Page, R. A., and Jones, P. L. (2014a). Risks of multimodal signaling: bat predators attend to dynamic motion in frog sexual displays. *Journal of Experimental Biology*, 217(17):3038–3044.
- Halfwerk, W., Jones, P. L., Taylor, R. C., Ryan, M. J., and Page, R. A. (2014b). Risky ripples allow bats and frogs to eavesdrop on a multisensory sexual display. *Science*, 343(6169):413–416.
- Harper, D. G. C. (2006). Maynard Smith: amplifying the reasons for signal reliability. *Journal of Theoretical Biology*, 239(2):203–9.
- Hauglund, K., Hagen, S. B., and Lampe, H. M. (2006). Responses of domestic chicks (*Gallus gallus domesticus*) to multimodal aposematic signals. *Behavioral Ecology*, 17(3):392–398.

- Haynes, K. F. and Yeargan, K. V. (1999). Exploitation of intraspecific communication systems: Illicit signalers and receivers. *Annals of the Entomological Society of America*, 92(6):960–970.
- Hebets, E. A. and Papaj, D. R. (2005). Complex signal function: developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology*, 57(3):197–214.
- Hebets, E. A., Vink, C. J., Sullivan-Beckers, L., and Rosenthal, M. F. (2013). The dominance of seismic signaling and selection for signal complexity in *Schizocosa* multimodal courtship displays. *Behavioral Ecology and Sociobiology*, 67(9):1483–1498.
- Hughes, M. (1996). The function of concurrent signals: visual and chemical communication in snapping shrimp. *Animal Behaviour*, 52(2):247–257.
- Iwasa, Y. and Pomiankowski, A. (1994). The evolution of mate preferences for multiple sexual ornaments. *Evolution*, 48(3):853–867.
- Jennions, M. D. and Petrie, M. (1997). Variation in mate choice and mating preferences: a review of causes and consequences. *Biological Reviews of the Cambridge Philosophical Society*, 72(2):283–327.
- Johnstone, R. A. (1995). Honest advertisement of multiple qualities using multiple signals. *Journal of Theoretical Biology*, 177(1):87–94.
- Johnstone, R. A. (1996). Multiple Displays in Animal Communication: ‘Backup Signals’ and ‘Multiple Messages’. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 351(1337):329–338.
- Johnstone, R. A. (1997). Recognition and the evolution of distinctive signatures: when does it pay to reveal identity? *Proceedings of the Royal Society B: Biological Sciences*, 264(1387):1547–1553.
- Kazemi, B., Gamberale-Stille, G., Tullberg, B. S., and Leimar, O. (2014). Stimulus salience as an explanation for imperfect mimicry. *Current Biology*, 24:965–969.
- Kelly, D. J. and Marples, N. M. (2004). The effects of novel odour and colour cues on food acceptance by the zebra finch, *Taeniopygia guttata*. *Animal Behaviour*, 68(5):1049–1054.

- Kilner, R. M. (1997). Mouth colour is a reliable signal of need in begging canary nestlings. *Proceedings of the Royal Society B: Biological Sciences*, 264(1384):963–968.
- Kim, S.-Y., Noguera, J. C., Morales, J., and Velando, A. (2011). The evolution of multi-component begging display in gull chicks: sibling competition and genetic variability. *Animal Behaviour*, 82(1):113–118.
- Kondo, N., Izawa, E.-I., and Watanabe, S. (2012). Crows cross-modally recognize group members but not non-group members. *Proceedings of the Royal Society B: Biological Sciences*, 279(1735):1937–42.
- Kulahci, I. G., Dornhaus, A., and Papaj, D. R. (2008). Multimodal signals enhance decision making in foraging bumble-bees. *Proceedings of the Royal Society B: Biological Sciences*, 275(1636):797–802.
- Leonard, A. S., Dornhaus, A., and Papaj, D. R. (2011). Flowers help bees cope with uncertainty: signal detection and the function of floral complexity. *The Journal of Experimental Biology*, 214(1):113–21.
- Leonard, M., Horn, A., and Parks, E. (2003). The role of posturing and calling in the begging display of nestling birds. *Behavioral Ecology and Sociobiology*, 54(2):188–193.
- Lohrey, A. K., Clark, D. L., Gordon, S. D., and Uetz, G. W. (2009). Antipredator responses of wolf spiders (Araneae: Lycosidae) to sensory cues representing an avian predator. *Animal Behaviour*, 77(4):813–821.
- Mackintosh, N. J. (1973). Stimulus selection: Learning to ignore stimuli that predict no change in reinforcement. In Hinde, R. and Stevenson-Hinde, J., editors, *Constraints on Learning: Limitations and Predispositions.*, page 488. Academic Press, Oxford.
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82(4):276–298.
- Magnhagen, C. (1991). Predation risk as a cost of reproduction. *Trends in Ecology & Evolution*, 6(6):183–6.

- Marples, N., van Veelen, W., and Brakefield, P. (1994). The relative importance of colour, taste and smell in the protection of an aposematic insect *Coccinella septempunctata*. *Animal Behaviour*, 48(4):967–974.
- Masters, W. M. (1979). Insect disturbance stridulation: its defensive role. *Behavioral Ecology and Sociobiology*, 5(2):187–200.
- McClintock, W. J. and Uetz, G. W. (1996). Female choice and pre-existing bias: visual cues during courtship in two *Schizocosa* wolf spiders (Araneae: Lycosidae). *Animal Behaviour*, 52(1):167–181.
- McLinn, C. M. and Stephens, D. W. (2006). What makes information valuable: signal reliability and environmental uncertainty. *Animal Behaviour*, 71(5):1119–1129.
- Milner, R. N. C., Jennions, M. D., and Backwell, P. R. Y. (2010). Eavesdropping in crabs: an agency for lady detection. *Biology Letters*, 6(6):755–757.
- Moller, A. and Pomiankowski, A. (1993). Why have birds got multiple sexual ornaments? *Behavioral Ecology and Sociobiology*, 32(3):167–176.
- Narins, P. and Grabul, D. (2005). Cross-modal integration in a dart-poison frog. *Proceedings of the National Academy of Sciences*, 102(7):2425–2429.
- Nieh, J. C., Barreto, L. S., Contrera, F. A. L., and Imperatriz-Fonseca, V. L. (2004). Olfactory eavesdropping by a competitively foraging stingless bee, *Trigona spinipes*. *Proceedings of the Royal Society B: Biological Sciences*, 271(1548):1633–1640.
- Olofsson, M., Jakobsson, S., and Wiklund, C. (2012). Auditory defence in the peacock butterfly (*Inachis io*) against mice (*Apodemus flavicollis* and *A. sylvaticus*). *Behavioral Ecology and Sociobiology*, 66(2):209–215.
- Owren, M. J., Rendall, D., and Ryan, M. J. (2010). Redefining animal signaling: influence versus information in communication. *Biology and Philosophy*, 25(5):755–780.
- Page, R. A. and Ryan, M. J. (2008). The effect of signal complexity on localization performance in bats that localize frog calls. *Animal Behaviour*, 76(3):761–769.

- Page, R. B. and Jaeger, R. G. (2004). Multimodal signals, imperfect information, and identification of sex in red-backed salamanders (*Plethodon cinereus*). *Behavioral Ecology and Sociobiology*, 56(2):132–139.
- Partan, S. R. (2013). Ten unanswered questions in multimodal communication. *Behavioral Ecology and Sociobiology*, 67(9):1523–1539.
- Partan, S. R. and Marler, P. (1999). Communication goes multimodal. *Science*, 283(5406):1272–1273.
- Partan, S. R. and Marler, P. (2005). Issues in the classification of multimodal communication signals. *The American Naturalist*, 166(2):231–45.
- Pavlov, I. (1927). *Conditioned Reflexes*. Oxford University Press, Oxford.
- Pearce, J. M. and Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, 52:111–139.
- Pearce, J. M. and Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87(6):532–552.
- Pearson, D. L. (1989). What is the adaptive significance of multicomponent defensive repertoires? *Oikos*, 54(2):251–253.
- Persons, M. H. and Uetz, G. W. (2005). Sexual cannibalism and mate choice decisions in wolf spiders: Influence of male size and secondary sexual characters. *Animal Behaviour*, 69(1):83–94.
- Pluta, S. R., Rowland, B. A., Stanford, T. R., and Stein, B. E. (2011). Alterations to multisensory and unisensory integration by stimulus competition. *Journal of Neurophysiology*, 106(6):3091–3101.
- Pomiankowski, A. and Iwasa, Y. (1993). Evolution of multiple sexual preferences by Fisher’s runaway process of sexual selection. *Proceedings of the Royal Society B: Biological Sciences*, 253(1337):173–181.

- Proops, L., McComb, K., and Reby, D. (2008). Cross-modal individual recognition in domestic horses (*Equus caballus*). *Proceedings of the National Academy of Sciences*, 106(3):1–5.
- Pruden, A. J. and Uetz, G. W. (2004). Assessment of potential predation costs of male decoration and courtship display in wolf spiders using video digitization and playback. *Journal of Insect Behavior*, 17(1):67–80.
- Prudic, K. L., Skemp, A. K., and Papaj, D. R. (2006). Aposematic coloration, luminance contrast, and the benefits of conspicuousness. *Behavioral Ecology*, 18(1):41–46.
- Ratcliffe, J. M. and Nydam, M. L. (2008). Multimodal warning signals for a multiple predator world. *Nature*, 455(7209):96–9.
- Rendall, D., Owren, M. J., and Ryan, M. J. (2009). What do animal signals mean? *Animal Behaviour*, 78(2):233–240.
- Rescorla, R. and Wagner, A. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In Black, A. and Prokasy, W., editors, *Classical Conditioning II: Current Research and Theory*, pages 64–99. Appleton Century Crofts, New York.
- Reznick, D. N. and Endler, J. A. (1982). The impact of predation on life history evolution in Trinidadian Gruppies (*Poecilia reticulata*). *Evolution*, 36(1):160–177.
- Rhebergen, F., Taylor, R. C., Ryan, M. J., Page, R. A., and Halfwerk, W. (2015). Multimodal cues improve prey localization under complex environmental conditions. *Proceedings of the Royal Society B: Biological Sciences*, 282(1814):20151403.
- Roberts, J. A., Taylor, P. W., and Uetz, G. W. (2006). Consequences of complex signaling: predator detection of multimodal cues. *Behavioral Ecology*, 18(1):236–240.
- Roberts, J. A. and Uetz, G. W. (2008). Discrimination of variation in a male signaling trait affects detection time in visual predators. *Ethology*, 114:557–563.
- Rowe, C. (1999). Receiver psychology and the evolution of multicomponent signals. *Animal Behaviour*, 58(5):921–931.

- Rowe, C. (2002). Sound improves visual discrimination learning in avian predators. *Proceedings of the Royal Society B: Biological Sciences*, 269(1498):1353–7.
- Rowe, C. and Guilford, T. (1999). The evolution of multimodal warning displays. *Evolutionary Ecology*, 13:655–672.
- Rowe, C. and Halpin, C. (2013). Why are warning displays multimodal? *Behavioral Ecology and Sociobiology*, 67(9):1425–1439.
- Rowe, C., Lindström, L., and Lyytinen, A. (2004). The importance of pattern similarity between Müllerian mimics in predator avoidance learning. *Proceedings of the Royal Society B: Biological Sciences*, 271(1537):407–13.
- Rowe, C. and Skelhorn, J. (2004). Avian psychology and communication. *Proceedings of the Royal Society B: Biological Sciences*, 271(1547):1435–42.
- Rubi, T. L. and Stephens, D. W. (2015). (in press) Should receivers follow multiple signal components? An economic perspective. *Behavioral Ecology*. doi: 10.1093/beheco/arv121.
- Rubi, T. L. and Stephens, D. W. (2016). (in revision) Why complex signals matter, sometimes. In Bee, M. A. and Miller, C. T., editors, *Perception and Cognition in Animal Communication*. Springer, New York.
- Ryan, M. J. and Cummings, M. E. (2013). Perceptual biases and mate choice. *Annual Review of Ecology, Evolution, and Systematics*, 44:437–459.
- Schluter, D. and Price, T. (1993). Honesty, perception and population divergence in sexually selected traits. *Proceedings of the Royal Society B: Biological Sciences*, 253(1336):117–22.
- Schmidt, V. and Schaefer, H. M. (2004). Unlearned preference for red may facilitate recognition of palatable food in young omnivorous birds. *Evolutionary Ecology Research*, 6(6):919–925.
- Siddall, E. C. and Marples, N. M. (2008). Better to be bimodal: the interaction of color and odor on learning and memory. *Behavioral Ecology*, 19(2):425–432.

- Skelhorn, J. and Rowe, C. (2006). Avian predators taste-reject aposematic prey on the basis of their chemical defence. *Biology Letters*, 2(3):348–350.
- Smith, C. L. and Evans, C. S. (2008). Multimodal signaling in fowl, *Gallus gallus*. *The Journal of Experimental Biology*, 211(13):2052–7.
- Stanislaw, H. and Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, 31(1):137–49.
- Stephens, D. W. and Dunlap, A. S. (2009). Why do animals make better choices in patch-leaving problems? *Behavioural Processes*, 80(3):252–60.
- Stratton, G. E. and Uetz, G. W. (1983). Communication via substratum-coupled stridulation and reproductive isolation in wolf spiders (Araneae: Lycosidae). *Animal Behaviour*, 31(1):164–172.
- Stratton, G. E. and Uetz, G. W. (1986). The inheritance of courtship behavior and its role as a reproductive isolating mechanism in two species of *Schizocosa* wolf spiders (Araneae; Lycosidae). *Evolution*, 40(1):129–141.
- Stynoski, J. L. and Noble, V. R. (2011). To beg or to freeze: multimodal sensory integration directs behavior in a tadpole. *Behavioral Ecology and Sociobiology*, 66(2):191–199.
- Tanner, W. P. J. and Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61(6):401–409.
- Taylor, R. and Ryan, M. (2013). Interactions of multisensory components perceptually rescue túngara frog mating signals. *Science*, 341(6143):18–20.
- Tibbetts, E. A. (2002). Visual signals of individual identity in the wasp *Polistes fuscatus*. *Proceedings of the Royal Society B: Biological Sciences*, 269(1499):1423–1428.
- Tibbetts, E. A. (2004). Complex social behaviour can select for variability in visual features: a case study in *Polistes* wasps. *Proceedings of the Royal Society B: Biological Sciences*, 271(1551):1955–1960.
- Tibbetts, E. A. and Dale, J. (2007). Individual recognition: it is good to be different. *Trends in Ecology and Evolution*, 22(10):529–537.

- Uetz, G. and Denterlein, G. (1979). Courtship behavior, habitat, and reproductive isolation in *Schizocosa rovneri* Uetz and Dondale (Araneae: Lycosidae). *Journal of Arachnology*, 7(2):121–128.
- Uetz, G. W., Clark, D. L., Roberts, J. A., and Rector, M. (2010). Effect of visual background complexity and light level on the detection of visual signals of male *Schizocosa ocreata* wolf spiders by female conspecifics. *Behavioral Ecology and Sociobiology*, 65(4):753–761.
- Uetz, G. W. and Norton, S. (2007). Preference for male traits in female wolf spiders varies with the choice of available males, female age and reproductive state. *Behavioral Ecology and Sociobiology*, 61(4):631–641.
- Uetz, G. W., Papke, R., and Kilinc, B. (2002). Influence of feeding regime on body size, body condition and a male secondary sexual character in *Schizocosa ocreata* wolf spiders (Araneae, Lycosidae): condition-dependence in a visual signaling trait. *Journal of Arachnology*, 30(3):461–469.
- Uetz, G. W., Roberts, J. A., and Taylor, P. W. (2009). Multimodal communication and mate choice in wolf spiders: female response to multimodal versus unimodal signals. *Animal Behaviour*, 78(2):299–305.
- Vallin, A., Jakobsson, S., Lind, J., and Wiklund, C. (2005). Prey survival by predator intimidation: an experimental study of peacock butterfly defence against blue tits. *Proceedings of the Royal Society B: Biological Sciences*, 272(1569):1203–1207.
- VanderSal, N. D. and Hebets, E. A. (2007). Cross-modal effects on learning: a seismic stimulus improves color discrimination learning in a jumping spider. *The Journal of Experimental Biology*, 210(20):3689–95.
- Wiley, R. H. (2006). Signal detection and animal communication. *Advances in the Study of Behavior*, 36:217–247.
- Wilson, A. J., Dean, M., and Higham, J. P. (2013). A game theoretic approach to multimodal communication. *Behavioral Ecology and Sociobiology*, 67(9):1399–1415.

- Yorzinski, J. L., Patricelli, G. L., Babcock, J. S., Pearson, J. M., and Platt, M. L. (2013). Through their eyes: selective attention in peahens during courtship. *The Journal of Experimental Biology*, 216(16):3035–46.
- Zahavi, A. (1993). The fallacy of conventional signalling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 340(1292):227–30.
- Zuk, M. and Kolluru, G. (1998). Exploitation of sexual signals by predators and parasitoids. *The Quarterly Review of Biology*, 73(4):415–438.