The Value of Artificial Stimuli in Behavioral Research: Making the Case for Egg Rejection Studies in Avian Brood Parasitism

Márk E. Hauber*, Lainga Tong*, Miklós Bánt†, Rebecca Croston*, Tomáš Grim‡, Geoffrey I. N. Waterhouse§, Matthew D. Shawkey¶, Andrew B. Barron** & Csaba Moskát††

* Department of Psychology, Hunter College and the Graduate Center of the City University of New York, New York, NY, USA
† Department of Evolutionary Zoology, MTA-DE 'Lendület' Behavioural Ecology Research Group, University of Debrecen, Debrecen, Hungary
‡ Laboratory of Ornithology and Department of Zoology, Palacky University, Olomouc, Czech Republic
§ School of Chemical Sciences, University of Auckland, Auckland, New Zealand
¶ Department of Biology and Integrated BioScience Program, University of Akron, Akron, OH, USA
** Department of Biological Sciences, Macquarie University, Sydney, NSW, Australia
†† MTA-ELTE-MTM Ecology Research Group, Hungarian Academy of Sciences, c/o Biological Institute, Eötvös Lóránd University and Hungarian Natural History Museum, Budapest, Hungary

Abstract

Experimentation is at the heart of classical and modern behavioral ecology research. The manipulation of natural cues allows us to establish causation between aspects of the environment, both internal and external to organisms, and their effects on animals’ behaviors. In recognition systems research, including the quest to understand the coevolution of sensory cues and decision rules underlying the rejection of foreign eggs by hosts of avian brood parasites, artificial stimuli have been used extensively, but not without controversy. In response to repeated criticism about the value of artificial stimuli, we describe four potential benefits of using them in egg recognition research, two each at the proximate and ultimate levels of analysis: (1) the standardization of stimuli for developmental studies and (2) the disassociation of correlated traits of egg phenotypes used for sensory discrimination, as well as (3) the estimation of the strength of selection on parasitic egg mimicry and (4) the establishment of the evolved limits of sensory and cognitive plasticity. We also highlight constraints of the artificial stimulus approach and provide a specific test of whether responses to artificial cues can accurately predict responses to natural cues. Artificial stimuli have a general value in ethological research beyond research in brood parasitism and may be especially critical in field studies involving the manipulation of a single parameter, where other, confounding variables are difficult or impossible to control experimentally or statistically.

Introduction

Over 50 years ago, Niko Tinbergen performed classic experiments to determine whether egg size and coloration affected avian parents’ choices to incubate eggs or remove broken eggshells and thereby reduce predation on nests (Tinbergen 1951; Tinbergen et al. 1962). By testing competing predictions, drawn from alternative hypotheses at the same level of analysis, Tinbergen endorsed and illustrated the value of the alternative hypothesis-testing framework for evolutionary and mechanistic studies of animal behavior in the wild. A critical component of these experimental approaches was the use of artificial stimuli that were inspired by natural forms, but they either mimicked or exaggerated aspects of those through the use of...
artificial materials (e.g., oversized model eggs, brighter painted colors, and artificially larger spots, than seen in natural eggs). In this way, the experiments limited and defined both the modality and the degree of variation within and among egg traits to best isolate those features that predictably elicited natural behaviors in wild animals.

Researchers have frequently and productively used painted model eggs, as well as dyed natural eggs (Fig. 1) in the search for the recognition cues used by hosts of brood parasitic birds to reject foreign eggs: well over 10,000 such egg rejection experiments have been completed (reviewed in Grim 2007). In most of these studies, a model or painted-over natural ‘parasitic egg’ is placed into an active nest and monitored for several days to determine whether the egg is accepted, pecked, or ejected, or the nest is abandoned (e.g., Davies & Brooke 1989; Antonov et al. 2009; Moskát et al. 2014a,b).

Yet, in recent years, both the value and the general applicability of evolutionary conclusions drawn about natural behaviors, through the use of artificial stimuli in egg rejection research, have been repeatedly questioned, for example, in both peer reviews of our manuscripts and at our conference presentations. Here, we define an artificial egg stimulus, as any material and pigment that is not taken directly from nature; according to this definition, a natural or model egg dyed blue with a human-manufactured paint, to resemble the avian perception of the immaculate egg of an American robin *Turdus migratorius*, is still an artificial stimulus (Croston & Hauber 2014). For example, Honza et al. (2007) used artificial dyes, to test the chromatic basis of foreign egg rejection by song thrush *T. philomelos*. Avian visual modeling (Avilés 2008) was then applied to the reflectance spectra of the artificial colors used, and combined with experimental rates of egg rejection to characterize, for the first time, the sensory-perceptual basis of egg recognition in birds (Cassey et al. 2008). Several studies followed these early works, including those using conspecific eggs to characterize the fine scale perceptual cues causing egg rejection behavior in other host species (e.g., Avilés et al. 2010; Spottiswoode & Stevens 2010; Stevens et al. 2013a), but some of these also included pointed criticism that experiments with artificial egg colors, and the resulting perceptual modeling, were not relevant to evolutionary and ecological studies of brood parasitism in natural contexts. Recently, we prepared a new manuscript inspired by Honza et al. (2007), and eventually published it (Bán et al. 2013) but during peer-review, we repeatedly encountered several incarnations of a knee-jerk reaction to our use of artificial stimuli to

Fig. 1: Examples of the use of artificial stimuli in egg recognition research: a model parasitic egg resembling the color of the common redstart-cuckoo *Cuculus canorus* host-race’s egg in the nest of the (a) great tit *Parus major* and (b) common redstart *Phoenicurus phoenicurus*; (c–e) host’s own egg(s) painted with added artificial spots in the nests of great reed warblers *Acrocephalus arundinaceus*. Photo credits: T. Grim and I. Zsoldos.
infer not only mechanistic but also evolutionary conclusions from the study. For example, one referee wrote: ‘… the problem remains that exaggerating natural variation (e.g., Tinbergen’s experiments) is a very different thing from presenting novel traits (both are useful, of course). Yes, the authors can use experiments like this to ask questions about the cognitive mechanisms that constrain evolution but, they don’t necessarily tell us what mechanisms these birds are using in nature because they do not mimic natural variation. Thus, the results remain mechanistic.’

While we appreciate the direct nature of comments like these, we disagree with the premise and conclusion about whether artificial stimuli can be used to explore evolutionary questions of brood parasites. Specifically, Tinbergen and his followers, including ourselves, clearly recognize(d) that the mechanisms themselves are an evolved phenotype that in turn influences the expression and outcome of selective pressures. Thus, mechanistic and evolutionary questions are never uninformative about each other (Tabórsky 2014). Furthermore, for evolutionary studies aimed at understanding the causes and consequences of natural variation of cues and responses, it is assumed and understood that extant variation is the result of evolutionary forces that have constrained it (Samaš et al. 2014). In turn, extending or exploring the phenotype’s variable space beyond the natural range is precisely what we need to do to probe how selection might be acting on novel traits.

Additionally, from an evolutionary perspective, whether a parasitic egg is rejected because it is recognized as an egg or a non-egg (e.g., detritus, flower petal) in the nest cup, is equivalent at the level of the fitness outcomes to responding to brood parasitism (i.e., egg rejection: beneficial; egg acceptance: costly). In other words, no matter how and what hosts perceive/interpret about the different objects (including eggs) that they see in the nest, the only thing that matters from an ultimate/evolutionary perspective is the resulting fitness of the host and the parasite. Conceptually, the same criticism can also apply for the use of a natural, non-mimetic cuckoo egg: It, too, may be rejected because the host considers it a piece of flower petal or other detritus fallen in the nest, or it may accepted it because its appearance is so different from the hosts’ own eggs so as to not be considered an egg, but instead an integral nest construction material. What and whether artificial (and natural) eggs placed into the nest are considered as ‘eggs’ is an empirical question that requires detailed and careful experimental analyses (reviewed by Guiqueno & Sealy 2012), but these questions should not be answered based on human (peer-reviewer’s) a priori interpretation of what a naturalistic stimulus should look like and what constitutes a ‘caricature of nature’. If anything, recent brood parasitism research has taught us that over the course of a handful of decades, hosts can evolve brand new egg coloration to evade the costs of accepting mimetic parasite eggs (Spottiswoode & Stevens 2012), thus what may be a caricature today, might be reality tomorrow.

To illustrate our argument in the context of the aims of ethological research, we highlight four potential benefits of the use of artificial colors in the study of avian egg rejection behaviors; critically, again, these benefits span both the ultimate and proximate levels of analysis (Tinbergen 1963). We also use published data to illustrate to fellow researchers, and to respond to critics, how to assess whether experiments with artificial stimuli may be used to interpret natural variation in host responses to natural stimuli. Finally, we welcome David Lahti’s (2015) commentary in this journal that complements well our commentary below.

Four Potential Benefits of Artificial Stimuli in Egg Rejection Research

The standardization of stimuli for developmental studies, with a focus on repeatability

Individual hosts of brood parasites may consistently reject or accept naturally laid parasitic eggs across repeated parasitism events, may switch from being acceptors to being rejecters, (or vice versa), or may vary their responses based on other ecological cues. For example, older oriental reed warblers Acrocephalus orientalis are more likely to reject common cuckoo Cuculus canorus eggs than are younger warblers (Lotem et al. 1992). To understand the ontogenetic basis of egg recognition and rejection, including its experience dependence, and the roles of learning and maternal effects, requires experimentation with a standardized set of stimuli across different time points of the host’s lifespan (Samaš et al. 2011; Grim et al. 2014; Moskát et al. 2014a, b). Because natural egg coloration changes within days of laying in the nest (Moreno et al. 2011), as well as in storage under controlled conditions (Cass et al. 2010), and natural nests may be difficult to find in a timely manner and the donor species may be a protected or otherwise vulnerable taxon, it is not always possible, and/or ethically justifiable, to use natural eggs as consistent stimuli for developmental studies, including the study of repeatability. For example, repeatability estimation
requires the use of identical stimulus across repeated experiments with the same individual; as any two natural eggs are never identical, the only way to test repeatability robustly is through the use of artificial models (for details see Grim et al. 2014).

Disassociation of correlated phenotypic traits of eggs used for sensory discrimination

Once it has been established from observational and experimental studies whether and to what extent hosts reject natural parasitic eggs, further use of natural eggs to understand the sensory basis of egg recognition is a heuristically limited approach (de la Colina et al. 2012). Natural stimuli often show limited variability overall in multidimensional trait space, but exhibit extensive covariation between specific traits (e.g., avian feather colors: Stoddard & Prum 2011); for example, eggs of brown-headed cowbirds *Molothrus ater*, that are always rejected by American robins, are always smaller in size, beige in background, and heavily maculated, compared with the larger and blue immaculate eggs of this host species (Friedmann 1929). Any of these differing egg traits, or their combinations, may be the possible recognition cue(s) for egg rejection, but these traits might simply be physiologically or structurally constrained to covary. Thus, using natural cowbird eggs exclusively as egg rejection stimuli prevents testing the relative contribution of size, color, and maculation in American robin’s egg recognition process (Rothstein 1982; Croston & Hauber 2014). Instead, using unnatural combinations of natural variation (e.g., small blue model eggs) generates novel (artificial) models that can critically aid the characterization of the proximate basis of the egg rejection cues used by hosts to eliminate parasitic eggs in the nest.

The estimation of the strength of selection on parasitic egg mimicry

The rejection of parasitic eggs by hosts represents a critical selective pressure in the coevolutionary arms race that drives parasites to evolve increasingly mimetic eggs, which required increasingly fine-tuned sensory systems to be detected by hosts (Davies 2000). This is because female parasites have nil fitness when their eggs are rejected and, thus, represent an evolutionary dead end. Yet, some parasites lay highly mimetic eggs, many of which are still rejected, whereas other parasites lay inaccurately or poorly mimetic eggs, most of which are accepted (Stoddard & Stevens 2011). To characterize comparative patterns of egg rejection behaviors, and to reconstruct the evolutionary trajectories of how rejection behaviors have changed with exposure to brood parasitism, requires a standardized metric of egg rejection responses (Grim et al. 2011). These, by definition, cannot be based on responses to natural parasitic eggs, because the coevolutionary hypothesis assumes a reciprocal and dynamic process between hosts and parasites, which will result in varying degrees of host–brood parasite egg mimicry across different systems (Igic et al. 2012). Instead, using a specific, variably rejected model egg color can provide a metric of egg rejection directly comparable across host populations and species. Accordingly, analyzing the responses of different species of common cuckoo hosts in Europe against the same artificial egg color revealed that more discriminating and rejecting hosts are parasitized by perceptually more mimetic parasite eggs (Stoddard & Stevens 2011).

Establishment of the evolved limits of sensory and cognitive plasticity

Presenting hosts exclusively with foreign eggs that are within the natural range of variation can also lead to incorrect conclusions about whether hosts recognize and reject foreign eggs. For example, some hosts, including the common redstart *Phoenicurus phoenicurus* accept virtually all naturally laid parasitic common cuckoo eggs (Rutila et al. 2006). When experimentally testing such a host’s egg discrimination ability by introducing natural parasite eggs laid elsewhere, model eggs painted to resemble them, (Fig. 1), or host eggs only partially dyed, this host accepts most of these foreign egg types, too (Rutila et al. 2002; Hauber et al. 2014). The results would then lead to the conclusion that egg rejection as a defense against parasites has not evolved in the redstart. However, there is a biologically critical, alternative functional explanation: that even if such hosts have evolved sensory mechanisms to recognize increasingly similar foreign eggs, their recognition mechanisms may be circumvented by the high accuracy of the coevolved mimicry of the parasitic egg’s appearance; in other words, the cuckoo eggs are such a good match of the redstart eggs that they cannot be discriminated and, thus, rejected by this host. This alternative hypothesis can be directly tested solely through the use of artificial eggs that deviate in a known direction from the phenotypic range of natural host and parasite eggs; using natural eggs of other species, or even conspecifics, would introduce both tractable (measured) and
intractable (unmeasured) sources of variation. Once the host’s ability to reject such non-matching eggs has been established, experimenters can move onto the use of better matching (more mimetic) eggs to meaningfully isolate more proximate drivers of egg rejection. Similarly, most of the grassland passerines that lay beige and spotted eggs accept all or nearly all beige and spotted cowbird-like eggs, but reject blue model eggs (Klippelstine & Sealy 2008). Importantly, in hosts that do reject non-mimetic eggs, the use of increasingly mimetic models is needed to establish the sensory thresholds of these discrimination abilities, and then to test whether these perceptual acceptance thresholds function adaptively, that is, to allow the rejection of foreign eggs to reduce the fitness costs of brood parasitism (e.g. Croston & Hauber 2014).

Responses to Artificial Stimuli can Predict Behaviors in Response to Natural Stimuli: The Case of Egg Rejection by a Brood Parasite Host

We recognize here that the use of artificial colors and/or materials can also be a severe constraint on the utility of these experiments in evolutionary interpretations of egg rejection data, for example, when using treatments which only change the color of the egg in spectral ranges not perceived by the subject (Avilés et al. 2006), or when model eggs are made from materials that cannot be pierced or grasped for successful ejection, despite repeated rejection attempts by hosts (Antonov et al. 2009). Nonetheless, to evaluate our specific claim that, contrary to our critics, experimentally induced behaviors in response to artificial stimuli can help to explain both causation and pattern in fitness-relevant responses to natural cues, we focused on our own published data (Bán et al. 2013; Moskát et al. 2014a,b). Specifically, we tested for a predictive relationship between the evolved behavior (egg ejection) and the artificial stimuli (dyed egg colors) at nests of the free-living great reed warblers Acrocephalus arundinaceus, an intermediate rejecting host species of the common cuckoo Cuculus canorus in central Hungary (Bán et al. 2013). From that study, we obtained host responses to experimental parasitism with a single foreign egg (host egg dyed with a highlighter pen of one of five colors, n = 12–16 nests) and contrasted them with egg rejection rates of a natural conspecific egg (moved a different host’s nest, n = 16; Bártol et al. 2002) and a natural parasite egg (a cuckoo egg moved from a parasitized to a non-parasitized nest, n = 13 nests; C. Moskát, unpublished data). We then calculated a stimulus metric that can be applied to both artificial and natural color stimuli: we measured avian visible spectral reflectance (300–700 nm) and used perceptual modeling to estimate chromatic contrast distances between natural host eggs’ background coloration and stimulus egg coloration (Moskát et al. 2014a).

Our data points did not include the limits of rejection probabilities (0%, 100%), and so we used a linear regression analysis between egg rejection rates and pairwise just noticeable differences (chromatic JNDS, n = 8 randomized egg pairs per color type; Fig. 2); the result showed a significantly positive relationship between perceivable chromatic contrasts and egg rejection rates ($R^2 = 0.29$, $F_{5,38} = 15.3$, $p = 0.0004$). When we also plotted the mean values of JNDS and experimentally induced rejection rates of single, natural conspecific eggs or single, natural parasitic eggs among the data points from these artificial colors, the natural eggs fell within the 95% confidence interval of the predicted means (Fig. 2); the combined model, including both artificial and natural eggs, was also significant ($R^2 = 0.32$, $F_{5,45} = 20.7$, $p < 0.0001$). The implication is that behavioral responses to natural stimuli are within the range predicted by variation in behavioral responses elicited by diverse artificial stimuli.

Conclusions

Conceptually, our arguments go far beyond studies on egg rejection by hosts of avian brood parasites, as similar dyeing treatments are also used for

![Fig. 2: The relationship between egg rejection rates of great reed warblers in response to experimentally introduced eggs, and avian perceivable chromatic distances (chromatic JNDS) between natural coloration of the host’s own eggs and the artificial coloration of artificially dyed natural eggs, as well as of natural conspecific and natural parasitic, common cuckoo eggs. The graph depicts the mean JND and the percent of rejection per egg type, the regression line (straight line), and its 95% confidence intervals (curved lines).](image-url)
experimental studies on nest predation (Weidinger 2001), nest mate recognition (Tibbetts 2002), and in many other experimental fields of animal ecology, evolution, and behavior (Ferrari et al. 2008). For example, artificial stimuli that fall far outside of range of natural stimuli proved to be useful in non-brood parasitism studies, including, camouflage (Stevens et al. 2013b) and sexual selection (Safran et al. 2010). Here, we argue that experimental studies with wild animals should not be classified a priori as strictly mechanistic and discarded as irrelevant to fitness, on the basis that manipulations involve artificial stimuli in quantity, in quality, or in both. Instead, artificial stimuli should be appreciated and utilized when these allow for the careful design, alteration, and delivery of exact cues and triggers that elicit fitness-relevant responses in freely behaving animals. This is especially relevant for studies in the wild, where other social and ecological cues and contexts are typically uncontrolled, and most also remain unmeasured. In turn, the possibility to design specific stimulus types that vary (only) along known trait dimension(s) remains the core strength of behavioral experimentation. Implementing diverse and yet standardized stimuli can be informative for both proximate, mechanistic questions about developmental and cognitive processes, and for ultimate, comparative analyses of predicted behavioral responses induced by these stimuli, and their consequences on fitness. However, we also recognize that there are limits to the use and utility of artificial stimuli in the study of evolutionary processes (see Lahti 2015 Debate article in this journal). To address these concerns empirically, we recommend (and illustrate above, see Fig. 2) the use of statistical checks to assess whether chosen stimuli, and/or the behavioral responses elicited by these, fit or predict the known range of responses elicited by natural stimuli.

Acknowledgements

This Perspectives manuscript in Ethology was not handled editorially by any of the co-authors. Our work here was motivated by the many spirited peer-reviewer comments surrounding the publication of various egg rejection studies that we have authored. We thank the editors for their invitation to contribute this article to Ethology. Brani Igić for help with spectral analyses, and Stephen Rothstein, Daniel Hanley, David Lahti, and the Shawkey laboratory’s members for comments. Funding was provided by the Raab Presidential Research Fellowship and the McNair Scholars programs at Hunter College (to LT), a Human Frontier Science Program’s Young Investigator Award (to MEH, TG, MDS, and GINW), the NESCent Working Group: Toward a unified evolutionary theory of decision-making in animals (AFOSR grant FA9550-13-1-0222) and the Hungarian National Science Fund (OTKA No. 83217 to CM).

Literature Cited


