**Female audiences compose male courtship displays in a lek-mating bird**

Liam Taylor, Nicholas Oakley, and David McDonald

## ABSTRACT

Animal behavior research usually assumes the characteristics of courtship displays, such as quality or complexity, are intrinsic traits of the organism performing the display. Here we provide evidence that the composition and complexity of courtship displays are determined by the female audience, rather than the male performer, in lek-mating Golden-winged Manakins (*Masius chrysopterus*). We used video observation to analyze 423 male courtship displays in three audience contexts: Solo displays with no audience (SOLO, n = 308), unsuccessful displays for a female audience (AUDI, n = 108), and rare, successful displays for females ending in copulation (COP, n = 13). Using entropy and a novel metric we call compressibility, we found the arrangement of behavioral elements is the most complex (i.e., most unpredictable and varied) in SOLO displays, less complex in AUDI displays, and least complex in COP displays. Using Jaro string distance, a method from record-linkage theory, we directly quantified similarities in display composition. Displays varied more by audience context than individual male performer, and COP displays were arranged most uniformly. Finally, we showed males perform different behaviors based on the location and behavior of their female audience. We found no evidence that males are intrinsically constrained or consistent in arranging their displays. Our results suggest the courtship displays of Golden-winged Manakins are dynamic, joint compositions by females and males, not controlled demonstrations of male traits. Future studies of sexual selection and mate choice must distinguish features of organisms (e.g., genetic traits or preferences) from features of interactions between organisms.

**KEYWORDS**

compression, entropy, intra-action, Jaro distance, lek, repertoire complexity, syntactic complexity

## **INTRODUCTION**

Studies of sexual selection and animal behavior traditionally ask how performer traits (e.g., a peacock’s tail) influence audience behavior (e.g., a peafowl’s decision to copulate; Darwin, 1871; Prum, 2017). Quantitative models of sexual selection demonstrate how audience mating preferences can in turn result in the evolution of elaborate, intrinsic performer traits, even when those traits are costly to develop and maintain (e.g., Fisher, 1930; Kirkpatrick, 1987; Kokko et al., 2003; Lande, 1981). Inspired by theory, a vast amount of empirical work in animals investigates how audiences judge the properties of performer traits during mate choice (Andersson, 1994; Rosenthal, 2017). Although recent research increasingly posits environmental, social, or age-related plasticity in female preferences and male traits (e.g., DuVal et al., 2023; Fox et al., 2019), this perspective still assumes preferences and traits are, at the moment of a courtship display, identifiable properties of audiences and performers, respectively.

However, some studies of animal courtship suggest we cannot understand mate choice strictly in terms of audience preferences for intrinsic performer traits. For example, male Satin Bowerbirds (*Ptilonorynchus violaceus*) decrease the intensity of their movement displays in response to female crouching signals (Patricelli et al., 2002), and male fiddler crabs (*Austruca perplexa*) change the frequency of claw-waving displays when females approach (How et al., 2008). In these species, a female cannot objectively judge movement intensity or claw-waving frequency as intrinsic male traits; those characteristics change with respect to the expression of her own preferences.

We ask whether female audiences or male performers are more important in determining the repertoire, complexity, and composition of courtship displays in a tropical, lek-mating bird, the Golden-winged Manakin (Passeriformes: Pipridae: *Masius chrysopterus*, monotypic in the genus and hereafter *Masius*). Male *Masius* perform acrobatic courtship displays to entice prospective female mates, thereafter providing no parental care (Prum & Johnson, 1987; Snow & Snow, 1992). We used video observations to characterize hundreds of *Masius* displays in three audience contexts: solo male displays (SOLO), unsuccessful displays for a female audience (AUDI), and successful displays for a female audience that ended in copulation (COP).

First, we show how display characteristics—duration, repertoire size, entropy, and compressibility (a measure of simplicity and repetition)—vary across audience contexts, with COP displays being the simplest and most repetitive. Second, we use a novel application of record-linkage theory (Jaro string distance; Fellegi & Sunter, 1969; Jaro, 1989) to show displays are more similar when performed by different males in the same audience context than when performed by the same male in a different audience context. Third, we show how simple female behaviors directly influence the behavioral elements in male displays. These results indicate *Masius* courtship displays are dynamic interactions between audiences and performers, rather than straightforward demonstrations of intrinsic performer traits under sexual selection.

## METHODS

### Study Site

We collected data in 2015-2016 (December, June-August) and 2017 (September-December) at the Milpe Bird Sanctuary of the Mindo Cloudforest Foundation, in northwestern Ecuador (~0°1’48”N, 78°57’12” W). Milpe comprises 100 ha of west slope Chocó-Andean forest (1,100 m elevation) with a network of maintained trails.

*Study Species*

Male *Masius* tend small sections of fallen, mossy, slanted logs (~20–60 cm2 surface) on which they perform gymnastic displays (Prum & Johnson, 1987). Display logs are often in aural, but rarely in visual, contact with one another as part of a broader, dispersed lek (Bradbury, 1981). As in many other manakin species, female *Masius* have an overall green plumage while males undergo a two-year process of plumage maturation from predefinitive (green) to definitive (black and golden-yellow) plumages over multiple years (Schaedler et al., 2021). Some older predefinitive-plumaged males can be identified by waxy nape feathers, a golden horn-like crest, or golden forehead feathers. Females are known to copulate only with definitive-plumage males (*pers. obs.*).

### Field Methods

We used mist-nets (6 or 12 m length, 30 mm mesh) to capture and mark individual birds. All *Masius* were banded with a numbered aluminum band and a unique combination of plastic color bands. Other species were released immediately. All field methods were approved by the University of Wyoming Institutional Animal Use and Care Committee (Protocol#20160602DM00242-02).

After regular in-person observations to identify *Masius* display logs, we monitored logs with video cameras. We prioritized video surveillance at logs with high manakin activity or female visitation, or at newly discovered logs. Individual cameras (Sony Handycam HDR-CX405 or Sony Handycam HDR-CX240, Sony Corp., Tokyo, Japan) were housed in weatherproof containers and powered with external batteries (12v motorcycle battery or 10,000 mAh GETIHU power bank, Shenzhen Top Star Industry Co. Ltd., China). Recording ended when the memory card reached capacity, the batteries died, or weather necessitated camera retrieval. We ran 4 to 6 cameras daily, one per log, with each camera recording approximately 7 hrs of video. Footage was flagged with a motion detection program (<https://github.com/ltaylor2/Masius_Movement>) using OpenCV v3.3 (Bradski, 2000) in Python v2.7. Motion clips were then manually verified and bookmarked for subsequent coding.

*Displays and Behavioral Elements*

We defined a display as a sequence of distinct behavioral elements during which time males were never absent for >60 s. We included only sequences that lasted >60 s and that included at least one of two core behavioral elements: “Audible log-approach dive” or “Side-to-side bow” (Prum & Johnson, 1987; Taylor et al., 2020).

*Masius* displays can involve multiple males and females both as performers and audience members (Prum & Johnson, 1987; Taylor et al., 2020) For this study, we excluded all displays featuring multiple males dancing during the same display (n = 26) and displays where the only audience was confirmed males (n = 4, including one copulation). All displaying males in our final dataset had definitive plumage.

We categorized the remaining 423 male displays into three contexts: SOLO, AUDI, and COP. SOLO displays were solo male performances, AUDI displays featured a female audience member at some point during the display but did not end in successful copulation, and COP displays ended in a successful copulation. For our main analyses, we ended COP displays at first copulation. A comparison of before- and after-copulation displays is given in Supplementary Material.

We coded display elements using Behavioral Observation Research Interactive Software (BORIS), an open-source event logging tool (Friard & Gamba, 2016). A total of 41 elements occurred in the raw BORIS logs (Table S1). We subsequently excluded tracking elements (e.g., “Start”), movement-based elements (e.g., “Male1 On Log”, “Female Movement”), female responses (e.g., “Female Tracking Male”), and male behaviors not directly involved in display dances (e.g. “Vocalization,” “Gardening”). Finally, we retained only one half of paired elements (e.g., merging “Side-to-side bow Left” and “Side-to-side bow Right” into “Side-to-side Bow”) and excluded “Attempted copulation” and “Copulation” elements. The 12 remaining, core behavioral elements—including pauses (“Zero”)—are described in Table 1.

### Display Characteristics

We calculated five metrics related to organization and complexity of *Masius* courtship displays: (A) Duration, in seconds; (B) Length—the number of total elements; (C) Repertoire size—the number of distinct elements; (D) Entropy; and (E) Compressibility. Duration was calculated from the raw data before behavioral filtering. Whereas repertoire size measures a display in terms of the number of behavioral elements (Catchpole, 1987; Nowicki & Searcy, 2004), entropy and compressibility measure syntax, i.e. the arrangement of a given repertoire of elements. First-order entropy (Shannon, 1948) is a common measure of syntax complexity in animal behavior (e.g., Da silva et al., 2000; Freeberg & Lucas, 2012; McCowan et al., 2002; Palmero et al., 2014; Vanderbilt et al., 2015). A low-entropy system is associated with predictable patterns, whereas a high-entropy system is associated with greater “surprise.” However, entropy is more technically a probabilistic property of a source (e.g., a system that produces a display) than a direct property of a message (e.g., a display itself). To characterize display syntax more directly, we thus also used compression methods from computer science, which condense data based on redundant elements or repeated motifs (Huffman, 1952; Ziv & Lempel, 1977). A display with a simple or repeated syntax will be highly compressible. High compressibility is only possible in low entropy systems (Shannon, 1948), meaning these two metrics are not independent of one another.

To calculate entropy and compressibility, we translated displays into strings, where each kind of behavioral element was a distinct character (Table 1). We computed first-order entropy of display strings using the *entropy* function in R package *acss* (Gauvrit et al., 2016). The resulting entropy values depend on the empirical frequency of each element in each individual display. We thus scaled each value by the maximum possible entropy for a display—given as log2[number of distinct elements]—for a final metric between 0 and 1 (Vanderbilt et al., 2015). We computed compressibility using a standard text compression algorithm (LZ77, with Huffman Coding) implemented with default parameters in the R package *brotli* (Ooms & Google, Inc., 2022). Our final compressibility metric was compression ratio—the length of uncompressed string divided by the length of compressed string.

We compared our five metrics across display contexts (SOLO, AUDI, COP) with linear models, using AUDI as context intercept and including male ID and observation month as discrete fixed effects. Unidentified males were assumed to be unique individuals. To address our small sample size of COP displays (n = 13), we compared COP metrics to a randomized distribution. Across each of 100,000 replicates, we randomly selected (without replacement) 13 displays from any context in our full dataset. We then compared the distribution of randomly-selected metrics to empirical COP metrics. All analyses were conducted with the tidyverse packages in Program R v4.2 (R Core Team, 2022; Wickham et al., 2019).

### Display Similarity: Male performer vs. Audience context

We investigated whether the syntax of displays varied more in terms of audience context (SOLO, AUDI, COP) or in terms of male individuality using Jaro string distances. Jaro (1989) developed an elegant method for matching census records from disparate sources. The algorithm analyses the number of transpositions and mismatches between two strings of data and assigns a distance between 0 (no matches) and 1 (complete match). Low Jaro distance indicates two displays have a similar syntax.

We calculated the Jaro distance between all pairs of display strings with the R package *stringdist* (van der Loo, 2014)*.* Using only displays from identified males, we calculated four sets of distances between displays across the three contexts (SOLO, AUDI, COP): (A) same-male/same-context, (B) same-male/different-context, (C) different-male/same-context, and (D) different-male/different context.

We used a randomization procedure to confirm two COP-related comparisons. First, we compared different-male/same-context Jaro distances from COP displays (n = 39 distances) to same-male/different-context distances from COP displays (n = 2,030). Second, we compared different-male/same-context distances from COP displays (n = 39) to different-male/same-context distances from AUDI (n = 3,289) plus SOLO (n = 29,648) displays. For both comparisons, we drew 100,000 random sets of 39 distances, without replacement, from the comparison dataset.

### Female Behavior

We asked how simple female behaviors directly influenced male behavioral elements during the composition of a courtship display. For each male behavioral element in the dataset, females were marked as on or off the display log. When on the display log, females were marked as either above or below the male (i.e., upslope or downslope). We quantified the behavioral elements performed by males when females were on/off the log and above/below the displaying male for both AUDI and COP displays. Elements were excluded when female location was not coded.

## RESULTS

### Displays and Behavioral Elements

Our final dataset includes 423 *Masius* displays across 15 display logs (1-109 displays per log), with 308 SOLO, 102 AUDI, and 13 COP displays. Across all displays, 406 were performed by 10 different, identified males (1 to 199 displays each, Table S2). All displays occurred between 24-Jun and 14-Jan, with 77% of displays (324/423) occurring between October and December (Fig. S1). The earliest AUDI display was 30-Jun and the earliest COP display was 26-Oct. The dataset featured 33 AUDI displays with one of 10 identified females (1 to 8 attendances each) and 4 COP displays with one of 3 identified females (1 to 2 copulations each; Table S3). All 13 COP displays were performed by one of 3 identified males (1 to 9 copulation each; Table S4).

By definition, every display in our dataset featured at least one Audible log-approach dive or Side-to-side bow. Representation of the remaining behaviors differed in terms of display context (Table 1, see Table S5 for raw frequencies). Silent log-approach dives were frequent across SOLO displays but present in only 1 AUDI display and 0 COP displays. To-and-fro flights were common in AUDI displays (84/102 displays) but relatively rare in SOLO (85/308) and COP (2/13). Neck twists were common in AUDI (97/102) and COP (9/13), but rare in SOLO (27/308). Head-down bows were present in most SOLO (289/308) and AUDI (89/102) displays, but absent from COP. However, there were Head-down bows and the rare Metronome element in 2 and 1 after-copulation displays, respectively (Supplementary Material). Pauses, Mixed behaviors, and Other behaviors were more common across SOLO displays than AUDI and COP displays (Table 1).

### Display Characteristics

All five display characteristics—duration, length, repertoire size, entropy, and compressibility—varied significantly across audience contexts (Fig. 1). In terms of display duration, AUDI displays (mean ± SD: 204 ± 128 s) were significantly longer than either SOLO displays (133 ± 63 s) or COP displays (125 ± 41 s; model estimate AUDI *vs*. SOLO: -82, *P* < 0.001; AUDI *vs.* COP: -79, *P* < 0.001; overall *P* < 0.001, adjusted *R2­* = 0.18; Fig. 1A). In terms of display length, AUDI displays (94 ± 70 elements) were significantly longer than both COP displays (62 ± 16 elements) and SOLO displays (23 ± 13 elements; AUDI *vs.* SOLO: -76, *P* < 0.001; AUDI *vs.* COP: -37, *P* < 0.001; overall *P* < 0.001, *R2­* = 0.42; Fig. 1B). COP displays had significantly smaller repertoires (3.2 ± 0.8 distinct elements) than AUDI displays (5.5 ± 1.3 distinct elements), the latter of which did not differ from SOLO displays (5.6 ± 1.2 distinct elements; AUDI *vs.* SOLO: 0.1, *P* = 0.32; AUDI *vs.* COP: -2.3, *P* < 0.001; overall *P* < 0.001, *R2­* = 0.15; Fig. 1C). Across 100,000 replicates of 13 randomly-drawn displays, mean repertoire size was never less than the empirical mean of the 13 COP displays (Fig. S2).

SOLO displays showed significantly higher entropy (0.87 ± 0.08) than did AUDI displays (0.70 ± 0.16), which in turn had significantly higher entropy than did COP displays (0.26 ± 0.16; AUDI *vs.* SOLO: 0.17, *P* < 0.001; AUDI *vs.* COP: -0.44, overall *P* < 0.001, *R2­* = 0.58; Fig. 1D). There was an identical pattern in compressibility, measured as the ratio of uncompressed to compressed display string length. SOLO display strings had significantly lower compression ratios (1.04 ± 0.31) than AUDI display strings (2.73 ± 1.14), which were significantly less compressible than COP display strings (3.65 ± 1.01; AUDI *vs.* SOLO: -1.77, *P* < 0.001; AUDI *vs.* COP: 0.87, *P* < 0.001; overall *P* < 0.001, *R2­* = 0.63; Fig. 1E). None of 100,000 random sets of 13 displays had lower mean entropy or higher mean compression ratio than the empirical set of 13 COP displays (Fig. S2).

There were additional significant effects of observation months and individual males in for some metrics (Supplementary Material). As dictated by theory (Shannon, 1948) entropy and compressibility were significantly correlated across our dataset (Fig. S3). Our choice of compression algorithm (LZ77 and Huffman Coding) and our final compression ratio metric created an overall positive correlation between display length and compressibility (Fig. S4).

COP displays had a characteristic syntax (Table S2). Ten of 13 COP displays, across all 3 copulating males, ended with long (34+ element) stretches of Side-to-side bows followed by an Audible log-approach dive. One display (ID#1455) was similar except for the insertion of a single Half-bow before the dive, while another (ID#1987) had a short sequence of two additional elements—To-and-fro flights and then a Neck twist—before the final bow and dive combination. The remaining, unusual display (ID#1824) was primarily stretches of Neck twists; nevertheless, it ended with a bow and dive combination (Table S2).

### Display Similarity: Male performer vs. Audience context

Using Jaro distance, displays in every context (SOLO, AUDI, COP) were on average more similar to displays in the same context than to displays in other contexts, regardless of the male performing the display (Fig. 2). Although increasing Jaro distance was positively correlated with differences in display length and repertoire size, these correlations explained only a small fraction of the variance (Fig. S5).

Displays performed by the same male in the same context were on average slightly more similar to one another (Jaro distance, mean ± SD: 0.40 ± 0.13) than displays by different males in the same context (0.41 ± 0.13). In contrast, displays in different contexts were on average less similar to one another whether performed by the same male (0.55 ± 0.13) or different males (0.57 ± 0.14). The mean Jaro distance for COP *vs.* different-male/same-context displays (n = 39 distances) was lower than the mean of all 100,000 random draws (each n = 39) of COP *vs.* same-male/different context distances (Fig. S6).

On average, COP displays performed different males were more similar to one another (mean ± SD: 0.14 ± 0.10) than either AUDI displays (0.36 ± 0.11) or SOLO displays (0.41 ± 0.13). The mean distance for COP *vs.* different-male/same-context distances (n = 39) was lower than the mean of all 100,000 random draws (each n = 39) of AUDI or SOLO *vs.* different-male/same-context displays (Fig. S6).

Average similarity comparisons were not absolute rules. Of the 13 COP displays, 5 had the lowest Jaro distance to COP displays by the same male, and 2 to COP displays by a different male. However, 4 COP displays were most similar to AUDI displays by the same male and 2 to AUDI displays by a different male.

### Female Behavior

By definition, there were no female behaviors during SOLO displays. Females attended the display log for a significantly greater proportion of COP displays (92.0 ± 13.3% male elements with female on log) than AUDI displays (53.5 ± 24.4%; *t* = -8.71, d.f. = 24.9, *P* < 0.001, two-tailed t-test). Counting only times when females were on the display log, they stood downslope of the male for a significantly greater proportion of COP displays (97.2 ± 10.0% male elements with female downslope) than AUDI displays (74.7 ± 23%; *t* = -6.29, d.f. = 32.3, *P* < 0.001, two-tailed t-test). Indeed, females were only upslope of males for a total of 14 elements in a single COP display (ID-1824).

Female behaviors changed the representation of male behavioral elements during display (Table 2). Males almost universally performed Side-to-side bows when females were present on the log and standing downslope, which was the dominant situation in COP displays. Males were more likely to perform Head-down bows or Neck twists when females were instead standing upslope, which was more common in AUDI displays. On the rare occasion a female was upslope of the male during the one unusual COP display (ID-1824), the male continued to perform Neck twists. When females were off the log, males performed a greater proportion of Neck twists, To-and-fro flights, and Switch behaviors (Table 2).

## DISCUSSION

Our first major finding is that male *Masius* courtship displays become steadily simpler as they progress from those performed by male without a female audience (SOLO), to those performed for females but not resulting in copulation (AUDI), to those resulting in a copulation (COP; Fig. 1). Successful displays had significantly smaller repertoires of behavioral elements than all other displays (Fig. 1C). Progressive simplicity is further captured by entropy, which estimates the predictability in the arrangement of behavioral elements in a display. Successful *Masius* displays had the most predictable arrangements followed by unsuccessful displays for an audience, with solo male displays being the least predictable (i.e., highest entropy; Fig. 1D). In parallel, we used a novel compression ratio metric to directly quantify simplicity and repetition in the arrangement of behavioral elements in male *Masius* displays. Successful displays were the most highly compressible, followed by unsuccessful and then solo displays (Fig. 1E). Longer displays on average had simpler arrangements as measured by compressibility (Fig. S4), but successful displays were both shorter than unsuccessful displays for audiences (Fig. 1B) and more compressible (Fig. 1E). As in many studies of tropical lekking birds, directly observing copulations was extraordinarily rare (n = 13 COP displays across only 3 males). However, the small repertoires, low entropy, and high compressibility of our pseudoreplicated COP data exceeded any of 100,000 mean values randomly drawn from the total dataset (Fig. S2). Further, multiple displays by copulating males across different audience contexts allowed our linear models to isolate the relative effects of audience context and male identity.

Such results are usually interpreted as follows: female *Masius* prefer simpler displays, only copulating when males perform successful displays. Males thus attempt to perform those simple displays when females are in attendance, and successful males are those who can perform the simplest displays to match female preferences. Indeed, numerous studies investigate how display characteristics such as repertoire size and syntax complexity drive patterns in female mate choice (Catchpole, 1987; Da silva et al., 2000; Freeberg & Lucas, 2012; McCowan et al., 2002; Nowicki & Searcy, 2004; Palmero et al., 2014). These perspectives posit our display characteristics—duration, length, repertoire, entropy, compressibility—are intrinsic male traits.

We argue the opposite: female behaviors, not male traits, determine the composition of male *Masius* courtship displays. We support this assertion with two lines of evidence. First, displays showed higher similarity when performed by a different male in the same audience context than when performed by the same male in a different context (Fig. 2). For example, solo displays by male A are, on average, arranged more like solo displays by male B than displays for a female audience by male A. In other words, our measurements of the composition of display (Fig. 1C-E) reflect context—whether a female is present, and whether the display ends in copulation, regardless of male identity—rather than the male performer (Vanderbilt et al., 2015). Second, we find evidence that female attendance behaviors directly determine the behavioral elements of males during a display. Whether or not a display ends in copulation, male *Masius* almost always perform the Side-to-side bow behavior when females attend downslope on the display log (Table 2). During unsuccessful displays, females were often upslope on the log, or off the log altogether. As a result, males are more likely to perform different elements, especially Neck twists, To-and-fro flights, and Position switches, during unsuccessful displays (Table 2). When females remain downslope on the log preceding copulation, the consequence is a characteristic syntax for successful male displays: a uniform and predictable series of Side-to-side bow behaviors followed by an Audible log-approach dive (Table 1, Table S4).

There is no evidence that males were constrained from performing the proper display. By definition, all displays in our dataset contained one of the two key elements in successful displays: Side-to-side bows and Audible log-approach dives (Table 1). Successful displays are shorter and simpler than unsuccessful displays (Figs. 1). Solo displays had the largest repertoires (Fig. 1C). Any male *Masius* who performs a solo or unsuccessful display can perform the simple arrangements observed for successful displays. We thus cannot interpret the difference between solo displays and displays for female audiences as decreased effort or optimization on the part of solo males (How et al., 2008), nor can we interpret successful displays as honest indicators of male vigor or quality (Andersson, 1994).

Current theories of evolution by sexual selection assume variation in preferences and traits are related to heritable, genetic loci independently expressed in audiences and performers, respectively (Kirkpatrick, 1987; Kokko et al., 2003; Lande, 1981; Prum, 2010). In contrast, our measurements of *Masius* displays suggest the expression of male traits (i.e., male behaviors during display) are mediated by the expression of female preferences (i.e., female behavior on the display log). Females cannot therefore assess male traits without altering the features of the male trait during the process of assessment itself (Barad, 2006; Prum, 2023).

In their study of male bowerbirds responding to female cues, Patricelli et al. (2002) propose male responsiveness to females could itself be an intrinsic performer trait subject to sexual selection. How do certain males skillfully modulate their performance to please their female audience? How do females reliably select among variation in male responsiveness, when they can only assess that responsiveness through the controlled expression of their own preferences? One promising outlet involves comparing three modes of displays: ornaments, songs, and dances. Ornaments, such as tail feathers, will remain relatively fixed with respect to audience behavior. Songs can vary with audience behavior, but audiences can hear songs from a distance without interfering. Dances, as shown here, can vary as audiences observe and interact with the display.

## LITERATURE CITED

Andersson, M. (1994). *Sexual selection*. Princeton University Press.

Barad, K. (2006). *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press. https://doi.org/10.1215/9780822388128

Bradbury, J. W. (1981). The evolution of leks. *Natural Selection and Social Behavior*, 138–169.

Bradski, G. (2000). *The OpenCV Library* [Computer software].

Catchpole, C. K. (1987). Bird song, sexual selection and female choice. *Trends in Ecology & Evolution*, *2*(4), 94–97. https://doi.org/10.1016/0169-5347(87)90165-0

Da silva, M. L., Piqueira, J. R. C., & Vielliard, J. M. E. (2000). Using Shannon Entropy on Measuring the Individual Variability in the Rufous-bellied Thrush Turdus rufiventris Vocal Communication. *Journal of Theoretical Biology*, *207*(1), 57–64. https://doi.org/10.1006/jtbi.2000.2155

Darwin, C. (1871). *The descent of man and selection in relation to sex* (Vol. 1). D. Appleton.

DuVal, E. H., Fitzpatrick, C. L., Hobson, E. A., & Servedio, M. R. (2023). Inferred Attractiveness: A generalized mechanism for sexual selection that can maintain variation in traits and preferences over time. *PLOS Biology*, *21*(10), e3002269. https://doi.org/10.1371/journal.pbio.3002269

Fellegi, I. P., & Sunter, A. B. (1969). A theory for record linkage. *Journal of the American Statistical Association*, *64*(328), 1183–1210.

Fisher, R. A. (1930). *The Genetical Theory Of Natural Selection*. Clarendon Press. http://archive.org/details/geneticaltheoryo031631mbp

Fox, R. J., Fromhage, L., & Jennions, M. D. (2019). Sexual selection, phenotypic plasticity and female reproductive output. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1768), 20180184. https://doi.org/10.1098/rstb.2018.0184

Freeberg, T. M., & Lucas, J. R. (2012). Information theoretical approaches to chick-a-dee calls of Carolina chickadees (Poecile carolinensis). *Journal of Comparative Psychology*, *126*(1), 68.

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open‐source event‐logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, *7*(11), 1325–1330. https://doi.org/10.1111/2041-210X.12584

Gauvrit, N., Singmann, H., Soler-Toscano, F., & Zenil, H. (2016). Algorithmic complexity for psychology: A user-friendly implementation of the coding theorem method. *Behavior Research Methods*, *48*(1), 314–329. https://doi.org/10.3758/s13428-015-0574-3

How, M. J., Hemmi, J. M., Zeil, J., & Peters, R. (2008). Claw waving display changes with receiver distance in fiddler crabs, Uca perplexa. *Animal Behaviour*, *75*(3), 1015–1022. https://doi.org/10.1016/j.anbehav.2007.09.004

Huffman, D. A. (1952). A Method for the Construction of Minimum-Redundancy Codes. *Proceedings of the IRE*, *40*(9), 1098–1101. https://doi.org/10.1109/JRPROC.1952.273898

Jaro, M. A. (1989). Advances in Record-Linkage Methodology as Applied to Matching the 1985 Census of Tampa, Florida. *Journal of the American Statistical Association*, *84*(406), 414–420. https://doi.org/10.1080/01621459.1989.10478785

Kirkpatrick, M. (1987). Sexual selection by female choice in polygynous animals. *Annual Review of Ecology and Systematics*, 43–70.

Kokko, H., Brooks, R., Jennions, M. D., & Morley, J. (2003). The evolution of mate choice and mating biases. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1515), 653–664. https://doi.org/10.1098/rspb.2002.2235

Lande, R. (1981). Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences*, *78*(6), 3721–3725.

McCowan, B., Doyle, L. R., & Hanser, S. F. (2002). Using information theory to assess the diversity, complexity, and development of communicative repertoires. *Journal of Comparative Psychology*, *116*(2), 166.

Nowicki, S., & Searcy, W. A. (2004). Song Function and the Evolution of Female Preferences: Why Birds Sing, Why Brains Matter. *Annals of the New York Academy of Sciences*, *1016*(1), 704–723. https://doi.org/10.1196/annals.1298.012

Ooms, J. & Google, Inc. (2022). *brotli: A Compression Format Optimized for the Web* (R package version 1.3.0) [Computer software]. https://CRAN.R-project.org/package=brotli

Palmero, A. M., Espelosín, J., Laiolo, P., & Illera, J. C. (2014). Information theory reveals that individual birds do not alter song complexity when varying song length. *Animal Behaviour*, *87*, 153–163.

Patricelli, G. L., Uy, J. A. C., Walsh, G., & Borgia, G. (2002). Male displays adjusted to female’s response. *Nature*, *415*(6869), Article 6869. https://doi.org/10.1038/415279a

Prum, R. O. (2010). The Lande–Kirkpatrick mechanism is the null model of evolution by intersexual selection: Implications for meaning, honesty, and design in intersexual signals. *Evolution*, *64*(11), 3085–3100.

Prum, R. O. (2017). *The evolution of beauty: How Darwin’s forgotten theory of mate choice shapes the animal world-and us*. Anchor.

Prum, R. O. (2023). *Performance All the Way Down: Genes, Development, and Sexual Difference*. University of Chicago Press. https://books.google.com/books?hl=en&lr=&id=3iXMEAAAQBAJ&oi=fnd&pg=PP6&dq=prum+performativity&ots=7tcHnfQ7r4&sig=i8e-gAHlF5eYdZXOJJrRjB\_Dsps

Prum, R. O., & Johnson, A. E. (1987). Display behavior, foraging ecology, and systematics of the Golden-winged Manakin (*Masius chrysopterus*). *Wilson Bulletin*, *87*.

R Core Team. (2022). *R: A language and environment for statistical computing* (4.2.2) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rosenthal, G. G. (2017). *Mate Choice: The Evolution of Sexual Decision Making from Microbes to Humans*. Princeton University Press.

Schaedler, L. M., Taylor, L. U., Prum, R. O., & Anciães, M. (2021). Constraint and Function in the Predefinitive Plumages of Manakins (Aves: Pipridae). *Integrative and Comparative Biology*, *61*(4), 1363–1377. https://doi.org/10.1093/icb/icab063

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Snow, D. W., & Snow, B. (1992). Display of the golden-winged manakin Masius chrysopterus. *Bulletin of the British Ornithological Club*, *112*, 264–270.

Taylor, L. U., Oakley, N., & McDonald, D. (2020). Golden-winged Manakin (*Masius chrysopterus*). In T. S. Schulenberg (Ed.), *Birds of the World* (1.0). Cornell Lab of Ornithology. https://neotropical.birds.cornell.edu/Species-Account/nb/species/gowman1/overview

van der Loo, M. P. J. (2014). The stringdist package for approximate string matching. *The R Journal*, *6*(1), 111–122.

Vanderbilt, C. C., Kelley, J. P., & DuVal, E. H. (2015). Variation in the performance of cross-contextual displays suggests selection on dual-male phenotypes in a lekking bird. *Animal Behaviour*, *107*, 213–219. https://doi.org/10.1016/j.anbehav.2015.06.023

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., & Hester, J. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

Ziv, J., & Lempel, A. (1977). A universal algorithm for sequential data compression. *IEEE Transactions on Information Theory*, *23*(3), 337–343. https://doi.org/10.1109/TIT.1977.1055714