**Multiple meanings of complexity in the courtship displays of a lek-mating birds**

**ABSTRACT**

In polygynous lek mating systems, males perform elaborate courtship displays for females. Studies of courtship displays, especially vocal displays in birds, have linked reproductive success to behavioral complexity. These studies often address a simple level of complexity—repertoire complexity—which focuses on e.g., the number of unique display elements in a display. More recent work has begun addressing another level—syntax complexity—which focuses on how organisms arrange the elements of their display. We analyze 353 displays of lek-mating Golden-winged Manakins, *Masius chrysopterus,* by coding their dances as strings of behavioral elements. In addition to the simple measure of repertoire complexity (number of unique elements in a display), we use methods from information science to calculate multiple measures of syntax complexity (entropy and compressibility). We find that solo male displays (SOLO) were the most complex, whereas unsuccessful displays performed for a female audience (AUDI) were of intermediate complexity, and successful displays ending in copulation (COP, n = 13) were the simplest in terms of both repertoire and syntax. Using Jaro string distance, a method from record-linkage theory, we find that COP displays were the most uniform and that displays varied more between context (SOLO, AUDI, COP) than between individuals. Females chose to copulate after simple displays building to a dramatic high-speed element (the audible log-approach dive). These results raise questions about the difference between song and dance displays, the role of audience participation, and a third level of complexity—element complexity—that reckons with the aesthetic content of behaviors themselves.

**KEYWORDS**

complexity, repertoire, syntax, courtship, display element, lossless compression, entropy, Jaro string distance, lek, manakin

## **INTRODUCTION**

## **METHODS**

### *Study Site*

We collected data in 2015-2016 (January, 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 logs (~20–60 cm2 surface) on which they perform elaborate gymnastic displays. 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 process of plumage maturation from predefinitive (green) to definitive (black, golden-yellow, and/or orange-red) plumages over multiple years (Taylor et al. 2020, 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 only known to copulate with definitive-plumage males (*pers. obs.*).

### *Field Methods*

We used mist-nets (6 or 12 m length, 30mm 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 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 snewly discovered logs. Individual cameras (Sony Handycam HDR-CX405 or Sony Handycam HDR-CX240, Sony Corp., Tokyo, Japan) were housed in weatherproof container and powered with an external batter (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 filtered with a liberal motion detection program ($GITHUB) using the OpenCV library ($CITE) in Python v. ($CITE). After being flagged by the motion detection program, motion clips were 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 only included sequences that lasted >60 s and included two core behavioral elements, “Audible log-approach dive” and “Side-to-side bow” (Table 1; $PRUM\_JOHNSON, $TAYLOR).

*Masius* displays can involve multiple males and females both as performers and audience members ($PRUM\_JOHNSON, Taylor et al. 2018). For this study, we excluded all displays featuring multiple dancing males (n = 26) or a single, identified male audience (n = 3, all predefinitive Male #980, including one copulation). We retained 21 displays where an audience member was suspected as a predefinitive male based on plumage but performed no display behavior. A comparison of displays for female and suspected predefinitive male audiences is given in Supplementary Material.

We categorized the remaining 353 male displays into three categories: SOLO (n = 251), AUDI (n = 89), and COP (n = 13). SOLO displays were solo male performances, AUDI displays featured one or two audience members 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 the moment of 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 data files (Table S1). We 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 combined paired elements into single behaviors (e.g., “Side-to-side bow Left” and “Side-to-side bow Right” become “Side-to-side bow”) and excluded “Attempted copulation” and “Copulation” elements. The 12 remaining, core display behaviors—including pauses (“Zero”)—are described in Table 1.

### *Repertoire complexity*

We calculated three simple metrics related to the repertoire of individual *Masius* displays: (A) Duration, in seconds; (B) Length, in number of total elements; and (C) Repertoire size, in umber of unique elements. Display duration was calculated from the raw data, and thus included the timing of some elements excluded from other behavioral analyses.

We compared repertoire complexity measures across context (SOLO, AUDI, COP) with ANOVA and Tukey’s HSD. To address our small sample size of COP displays (n = 13), we compared our COP metrics to a randomized distribution. Across each of 10,000 replicates, we randomly selected (without replacement) 13 displays from across SOLO, AUDI, and COP displays. We then compared the distribution of the random replicate metrics to the values calculated using the full sample. All analyses were conducted with the *tidyverse* packages in Program R v4.2.2 ($TIDYVERSE, R).

*Syntax complexity*

We quantified syntax complexity of displays with two related measures: entropy and compressibility. We began by translating displays into strings of individual characters (Table 1). We computed the first-order entropy display strings with *entropy* function in R package *acss* ($CITE). The resulting values depend on the empirical frequency of each element in each individual display. We thus scaled each value by the maximum possible entropy for a display—given as log2(number of unique elements)—giving a final metric between 0 and 1 (Vanderbilt, Kelley, & DuVal 2015). We computed compressibility using a standard text compression algorithm (LZ77 and Huffman Coding) implemented with default parameters in the R package *brotli* ($CITE). The final compressibility metric was taken as the ratio of the length of the uncompressed string to the compressed string.

We compared syntax complexity measures across context (SOLO, AUDI, COP) with ANOVA and Tukey’s HSD. As with repertoire complexity, we used a randomization procedure to compare the small sample of 13 COP displays to a randomized distribution of both entropy and compressibility metrics (10,000 random sets of 13 displays, drawn with no replacement).

Entropy and compressibility are fundamentally intertwined metrics. In theory, high entropy systems produce outputs that lack well-defined, repeated motifs and are therefore difficult to compress ($CITE). We assessed the empirical correlation between entropy and compressibility in our data and investigated displays diverging from that correlation.

*Context vs. individual variation*

We asked whether displays varied more in terms of context (i.e., SOLO *vs.* AUDI *vs.* COP) or in terms of male individuality using Jaro string distances. Jaro (1989) developed a simple but elegant method for matching census records from disparate sources. The algorithm analyses the number of transpositions and mismatches between two strings and assigns a distance between 0 (no matches) and 1 (complete match).

We calculated Jaro distance among all display strings using R package *stringdist* ($CITE). For each context, we then compared three sets of distances: same-context/same-individual, same-context/different-individual, different-context/same-individual. Note we could only use displays from identified males for these comparisons.

## **RESULTS**

*Displays and behavioral elements*

Our final dataset includes 353 *Masius* displays across 15 display logs (1-102 displays per log). Of these displays, 341 were performed by one of 10 identified males (2-172 displays each, Table S2). All displays occurred between 24-Jun and 14-Jan, with the bulk of displays (278/353) between October and December (Fig. S1). The earliest AUDI display was 30-Jun and the earliest COP display was 26-Oct.

The dataset featured 36 AUDI displays with one of 11 identified females (1-7 attendances each) and 5 COP displays with one of 3 identified females (1-2 copulations each; Table S3). All 13 COP displays were performed by one of 3 identified males (1-9 copulation each; Table S4).

By definition, every display in our dataset featured at least one Audible log-approach dive and 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 (77/89 displays) but relatively rare in SOLO (63/251) and COP (2/13). Neck twists were common in AUDI (86/89) and COP (9/13), but rare in SOLO (21/251). Head-down bows were nearly universal in SOLO (248/251) and AUDI (88/89) 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, and Other behaviors were more common across SOLO displays than AUDI and COP displays (Table 1).

*Repertoire complexity*

In terms of raw display duration, AUDI displays (mean ± S.D.: 217 ± 130 s) were significantly longer than both SOLO displays (136 ± 65 s) and COP displays (126 ± 41 s; overall ANOVA *P* < 0.001; Fig. 1A). In terms of the total number of display elements, AUDI displays (101 ± 72 elements) were again significantly longer and more variable than SOLO display (62 ± 16 elements), which in turn were significantly longer than COP displays (24 ± 13 elements; overall ANOVA *P* < 0.001; Fig. 1B).

In contrast to these patterns in display length, COP displays had significantly smaller repertoires (3.2 ± 0.8 unique elements) than the similar SOLO displays (5.9 ± 1.1 unique elements) and AUDI displays (5.9 ± 1.0 unique elements; overall ANOVA *P* < 0.001; Fig. 1C). Across 10,000 replicates of 13 randomly-drawn displays from our dataset, mean repertoire size was never less than the empirical mean of the 13 COP displays (Fig. S2).

*Syntax complexity*

SOLO displays showed significantly higher scaled entropy values (0.87 ± 0.07) than AUDI displays (0.71 ± 0.12), which in turn had significantly higher scaled entropy than COP displays (0.26 ± 0.16; overall ANOVA *P* < 0.001; Fig. 2A). 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 (1.07 ± 0.30) than AUDI display strings (2.80 ± 1.15), which were significantly less compressible than COP display strings (3.65 ± 1.01; overall ANOVA *P* < 0.001; Fig. 2B). None of 10,000 random sets of 13 displays for each metric had a lower mean entropy or higher mean compression ratio than the empirical set of 13 COP displays (Fig. S2).

As expected, entropy and compressibility were significantly correlated across our dataset (linear regression, adjusted *R2 =* 0.53, *P* < 0.001;Fig. 2C). However, wide variation highlighted the differences between these metrics. For example, the most compressible display string (AUDI ID-1487, compression ratio = 7.45) had intermediate entropy (0.63) but was long, with 365 total elements, and primarily made up of long stretches of Side-to-side bows and Neck twists. A display with similar entropy but much lower compression ratio (AUDI ID-453, scaled entropy = 0.64, compression ratio = 2.53) was shorter, with 81 elements, and featured a tail of individual behaviors (coded “*IEDBA*”; Table 1) that made it difficult to compress. Indeed, 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. S3).

The lowest entropy display (COP ID-1533, scaled entropy = 0.11, compression ratio = 4.71) was simply 66 Side-to-side bows followed by an Audible log-approach dive. A closer look revealed a characteristic syntax for COP displays (Table S2). Ten of 13 COP displays, across all 3 copulating males, ended with long (34+) 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 display (ID-1824) was primarily stretches of Neck twists, yet still ended with a bow and dive combination (Table S2).

*Context vs. individual variation*

**DISCUSSION**

**LITERATURE CITED**

Figure 5 shows expected Jaro string distances (left four bars) under the hypothesis that individual differences between different males ("personalities") drive variation in display sequences. Low Jaro distances mean the displays compared are similar. The four gray and black bars on the right show the observed Jaro distances, which support the hypothesis that display sequences according to context (Mal, Fem, Cop), and do not support the individual-difference hypothesis depicted by the four leftmost expected bars. The pooled Within- and Between Mal Jaro distances differed significantly from the pooled Within- and Between Cop Jaro distances (*t-test*: *t* = 17.24, *df* = 91.2 P < 0.0001). To examine the effect of the small sample size (43) of Cop Jaro distance pairs between males, we randomly selected sets of 43 distance pairs and computed their mean Jaro distances from one another for 10,000 replicates. The absolute differences between the means of the resampled distributions and the observed means were < 0.0002, and within 0.1% of each other. None of the 10,000 random Mal-Within replicates (mean = 0.39; 95% CI = 0.36 to 0.43) overlapped the random Cop-Within replicates (mean = 0.25, 95% CI = 0.22 to 0.28). Likewise, no Mal-Between values (mean = 0.42; 95% CI = 0.39 to 0.45) overlapped with the Cop-Between values (mean = 0.23; 95% CI = 0.21 to 0.26). Differences between the Jaro distances across the three contexts (Mal, Fem, Cop) are shown in Fig. S2.

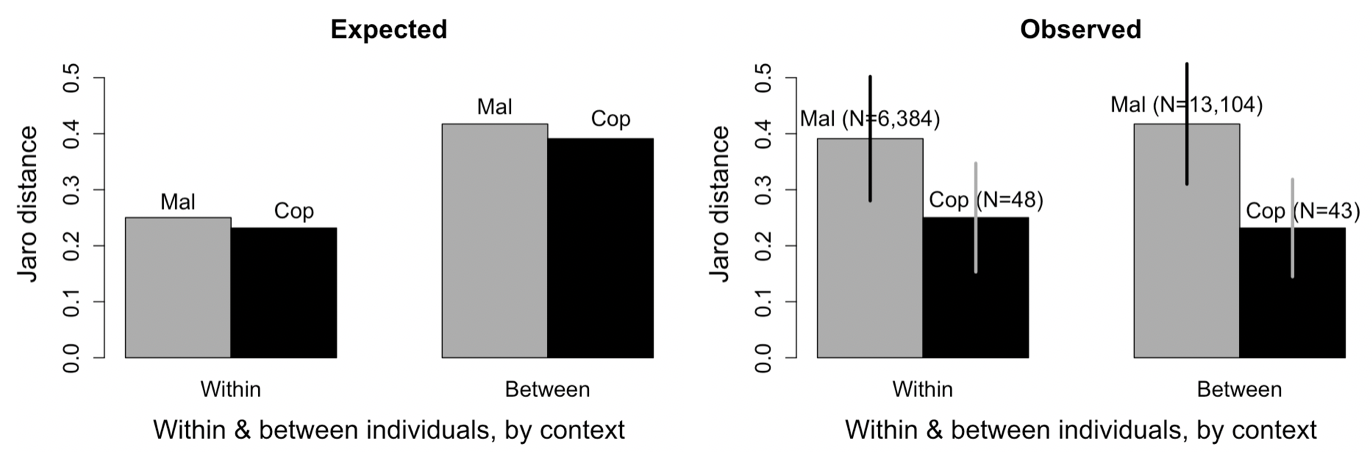


Figure 5. Expected (left four bars) Jaro string distance outcomes under the hypothesis that individual differences ("personalities") drive variation in display behaviour more than do contexts (Mal, Fem, Cop). If "personalities" mattered most, we would expect a male's own displays to be similar to each other (short Within Jaro distances; Mal-Within Cop-Within), regardless of context (Mal, Cop). We would expect a male's displays to be dissimilar to those of other males (high Mal-Between and Cop-Between). Instead (Observed, right four bars), a male's own Mal displays were almost as different from each other (Mal-Within = 0.39) as from those of other males (Mal-Between = 0.42). Cop displays were always very similar (short Jaro distances) whether Within or Between. In fact, males' Cop displays were slightly more similar (Cop-Between = 0.23) to those of other males, than to their own (Cop-Within = 0.25). Thus, variation in display is largely a matter of context; Mal Within resembles Mal Between and both differ greatly from Cop Between and Cop Within, which closely resemble each other, regardless of male or female identity. Vertical lines on rightmost four bars are ± 1 sd. None of the values in the Observed Mal-Within resampled distribution overlapped with any in the Observed Cop-Within resampled distribution. Likewise none of the values in the Observed Mal-Between resampled distribution overlapped with any of those in the Observed Cop-Between resampled distribution (see Methods and text for more details), meaning that the small sample size of Cop Jaro distances did not affect the conclusions.

## **DISCUSSION**

## Our major result is that context dependence and simplicity, not individual differences ("personalities") or complexity characterize variation in the courtship displays of male *Masius*. The pattern holds, across our three contexts (Mal, Fem, Cop), whether assessed in terms of weighted networks (Fig. 1), entropy (Fig. 2), compression ratio (Figs. 3 and 4) or Jaro string distance (Fig. 5). The emergent feature is a progression, from diverse (more random) display sequences that were dissimilar (from one bout to the next) when no female was present (Mal bouts), to redundant (predictable, organized) and similar to one another (within or between males) in displays leading to copulations (Cop bouts). In contrast, variation between males was relatively minor, especially in displays that included a copulation. Each of our analyses (networks, entropy, compression ratio, Jaro distances) contributes to a fuller understanding of how simplicity and context-dependence characterize the organization of *Masius* courtship displays.

## We assessed the comparative simplicity of the displays by analyzing them as networks (Fig. 1). Network density (proportion of all possible edges actually occurring) decreased from Mal to Fem to Cop, meaning that Mal bouts (density = 0.8) had many different kinds of transitions between display elements, while Cop bouts, especially (density = 0.27), had relatively few types of transitions. Low effective degree corresponds to high variance in edge weight, with one or a few edges dominating the transitions. With no variance in edge weights, effective degree equals simple binary degree (i.e., a simple count of the edges, disregarding weights). Cop bouts had low effective degree (2.8), meaning that most of the edge weight was concentrated on a few edges, particularly the repetition (heavy self-loops in Fig. 1) of Bows (before audible log-approach displays, ALADs) and Neck twists (after ALADs). In contrast, the Mal network had much higher effective degree (24.8), meaning a more even distribution of edge weights, and lower rates of repetition for any particular display element. All-simple-paths analysis showed a dramatic difference in the number of possible paths among nodes when comparing the Cop network (1,530 simple paths) to the Mal network (mean of 649,610 paths). That difference certainly underestimates the true difference by a wide margin, because we used a set of reduced-edge Mal networks to make the analysis computationally feasible.

## Commonly, increased song complexity, often assessed by entropy (Da Silva, Piqueria, & Vielliard, 2000; Palmero, Espelosín, Laiolo, & Illera, 2014) has been considered a positive attribute, associated with increased mating success. Here, we find that, to the contrary, Cop (successful courtship) bouts had significantly lower entropy (Fig. 2, mean = 1.57) than did Fem bouts (mean = 2.21) or Mal bouts (mean = 2.61). Thus, display bouts that include a copulation tended to be more predictable, with greater repetition of display elements, particularly Bows and Neck twists than do either Fem or Mal bouts. Concordant with our results, Vanderbilt, Kelley and DuVal (2015) found that in the dual-male courtship displays of Lance-tailed Manakins, *Chiroxiphia lanceolata*, displays for females had lower entropy than did displays in the absence of females.

## Compression ratio (Fig. 3) increased dramatically from Mal bouts to Fem bouts and thence to Cop bouts. Compression ratio was highest in Cop bouts, which featured long strings of repeated behaviours—particularly Bows and Neck twists—and the repeated motifs of elements detected by the LZ77 algorithm.

## Overall, Jaro string analysis (Fig. 5) indicated that Mal display bouts were variable within and between males, and that Cop bouts were similar, whether within or between males. Cop bouts (mean Jaro distance = 0.24) were more similar to each other than Fem bouts were to each other (mean Jaro distance = 0.38) or Mal bouts were to each other (mean Jaro distance = 0.41; the bars in Fig. S2). Thus, structural differences in the organization of courtship displays were due largely to the context—female present or not, and progressing toward copulation or not— and not to consistent differences between males in how they organized their displays. Although they measured similarity differently, Botero et al. (2009) stressed the importance of consistent song repertoires in Tropical Mockingbirds (*Mimus gilvus*), both as an ontogenetic trajectory and to explain reproductive success. Similarly, Janisch, Perinot & Fusani (2020) found that male *Manacus vitellinus* developed consistent courtship display sequences, in terms of movements around their courts.

## Why are Mal bouts (no female present) more random and disorganized than either Fem or Cop bouts? Displays with no female present (Mal bouts) could represent practicing, identifying the best trajectory for log-approach displays (ALADs and, particularly, SLADs), or maintaining control of a display log, to name but a few possibilities. None of these behaviours requires strict adherence to a fixed routine, possibly increasing the complexity and diversity of male-only displays (Mal) when compared to Fem or Cop bouts. The simple presence of a female appeared to suffice in galvanizing males to perform simpler and more organized display sequences.

## A feature that still puzzles us is how females choose which display logs to visit. In many lek-mating species, females can use fairly long-distance visual or acoustic cues to locate the courts of displaying males. In *Masius*, the "nurrt" advertising call is of surprisingly low-amplitude and is performed at a low rate overall, even at the height of display activity in Nov and Dec (2.2 nurrts/min at the two most active display logs) and within-bout vocalizations averaged only 4.2/min, reasonably similar to the calling rates documented by Prum & Johnson (1989). The call rate for *Masius* contrasts sharply with the far-carrying dual-male *toledos* of male Long-tailed Manakins, *Chiroxiphia linearis,* which occurred at overall rates > 5.4 toledos/min (McDonald, 1989) over the course of entire days (McDonald 2010, Fig. 1), with within-bout rates averaging 15.6/min (Trainer and McDonald, 1993). We conclude, provisionally, that males rely mostly on regular traffic of females to promote female visitation, as might be predicted under a hotspot model of lek placement (Bradbury, 1981; Beehler & Foster 1988). Further, our proportion of banded females and revisitation rates by banded females were far higher than in Long-tailed Manakins (pers. obs.), suggesting that female *Masius* may have much smaller home ranges that largely overlap the display logs of the males they sample.

## Our results do provide some insights into the bases and patterns of female mate choice, once they have chosen to visit a display log. The similarity of Cop displays across three distinct males, despite their being performed for at least 14 different females, suggests that a simple, rather invariant display worked well when females were ready to copulate, as indicated by females' steady presence on the display log during Cop bouts as compared to Fem bouts. A simple sequence of many Bows, followed by an ALAD, immediately followed by a copulation and then a series of Neck twists, was the consistent pattern for Cop bouts. One might suppose, that once having succeeded in copulating, males would modify their Mal and Fem displays to more closely resemble the simple rhythm of Cop displays. Our data do not support that hypothesis. Fig. 4 shows that M296's Mal displays did not become more organized and less complex as the breeding season progressed. Even immediately after Cop bouts with high compression ratios, M296's Fem and Mal bouts had low compression ratios was similar to those of bouts long before the first copulation.

## Once females chose to visit, an obvious feature of Fem and Cop displays was that male behaviour depended on the relative position of the female. When she was below him on the log, he would perform Bows until she moved or he left the log to perform an audible log-approach display (ALAD). When she was above him, or off the display log, he would perform Neck twists. If a female was not in the immediate vicinity of the display log, the male might still perform those same behaviours, but would tend not to repeat them for as long, and would perform them in broken series, in a less predictable order. Further, the male would intersperse other display elements that he did not perform while the female was on the log, such as Metronome displays. The result was long strings of disjointed behaviours and frequent male flights to and from the display log. Although it was not the primary focus of this paper, we found that female as well as male behaviour varied across contexts. Females moved on and off the display log more during Fem bouts, perhaps because it was their first visit to a particular display log that season, or because they lacked experience in evaluating males. The results suggest that when females were prepared to copulate, they moved much less on the log, providing some feedback to the male on the female's interest level, and causing males to settle into the simple and predictable rhythm of the peri-copulatory displays.

## The most dramatic behavioural element in the courtship display of *Masius* is the audible log-approach display (ALAD); it occurred just prior to all 24 observed copulations, and is perceived as a blur by the human eye (see slow-motion video at http://www.uwyo.edu/dbmcd/masiusvideo.html). Why should females respond to such an odd and rapid maneuver? One intriguing possibility is that it requires extreme agility that could be useful in a foraging context, for females as well as their offspring. Howe (1979) argued that small frugivores might be particularly vulnerable to predation at fruit trees, which are reliable sites for prey detection. Consistent with Howe's hypothesis, Uriarte et al. (2011) found that four species of manakins spent a mean of only 13 ± 4.3 secs at their fruit trees. Many of the visits had zero perch time, because birds often gulped (Levey, 1987) the fruits and departed, without ever perching. For larger frugivores, perch times can be considerably longer (median visit length 225 secs, Wheelwright, 1991). Ultra-rapid maneuvers while taking fruit might significantly reduce predation risk at these predictable locations. High-speed video of *Masius* at fruit trees, particularly in the event of a predation attempt, would be an interesting test of this hypothesis.

## Although we found that greater simplicity and uniformity characterized successful courtship displays, we are not arguing that simplicity or uniformity, per se, increases reproductive success in *Masius*. Instead, we conclude that context (presence or absence of a female), and behavioural feedback from female visitors, as demonstrated in the elegant experiment by Patricelli, Coleman & Borgia (2006; see also Rodríguez et al., 2012), together create a progression from diverse (unpatterned, with low repetition of elements), to orderly (highly patterned with repeated elements and motifs) in displays leading to copulations. Displays leading to successful copulations are simple and uniform from one iteration to the next, whether within the Cop displays of an individual male or between the Cop displays of different males. What remains an intriguing question is how and why females transition from fidgety (many landings and departures from the log in Fem bouts) to focused (steadily remaining on the log during Cop bouts). Gibson (1996) found that the inter-pop interval was positively correlated with the number of females visiting a sage grouse (*Centrocercus urophasianus*) male, but that eventual copulation was related to display rate. Similarly, McDonald (1989) found that female visitation correlated with the output of dual-male unison *toledo* song in Long-tailed Manakins. Copulatory success, however, correlated with the duration of the dual-male butterfly flights, which occur between bouts of dual-male backwards leapfrog displays. Those studies suggest that female preferences are based, not on a single male trait, but instead vary with a female's sampling context, experience and, very likely, her reproductive state. Thus, understanding prior female experience (both within and across years) may be critical to understanding how female response to courtship drives context-dependent variation in displays. The ball is back in the court of the difficult problem of assessing female sampling histories, often stretching back several years, and of ascertaining female motivational states.

*Conclusions*

## We conclude that *Masius* courtship displays progress from disorderly jumbles to tightly organized sequences characterized by repetition of major display elements (particularly Bows) as the context moves from male-only displays to those leading to copulations. We have provided a worked example of the utility of network, lossless compression and Jaro string-distance metrics that could be applied to many other sorts of sequences of behaviours. For example, the methods used here could be applied to the frame of reference, in terms of motor activity (sensu Miller, 1988), variation in song repertoires, other sorts of courtship display, animal movement trajectories and even sequences of environmental variables such as time series of temperature data. These metrics were powerful enough to differentiate between Fem bouts (female present) and Cop bouts (including one or more copulations) and to demonstrate that those differences arose due to differences in context, and not to individual differences between males. Despite the lack of individual variation in Cop displays, we do not conclude that behavioural syndromes are absent in *Masius*. They are just not a driving force behind the organization of and variation in courtship displays. Likewise, the simplicity of successful courtship displays (Cop bouts) does not mean that complexity and non-linearity (Bradbury & Vehrencamp, 2014; Patricelli & Hebets, 2016) are not important to many other aspects of the behaviour of *Masius* or other manakins. Nevertheless, for *Masius* displays, as found by Cardoso and Hu (2011) for wood warbler song, we conclude that "simple does it."

**DATA AVAILABILITY**

All R scripts used for the analyses are available, on request, from the corresponding author. [OR AS SUPPLEMENTARY MATERIAL?]

## **LITERATURE CITED**

Barske, J., Schlinger, B. A., Wikelski, M., & Fusani, L. (2011). Female choice for male motor skills. *Proceedings of the Royal Society of London Series B: Biological Sciences*, *278*, 3523–3528.

Beehler, B. M., & Foster, M. S. (1988). Hotshots, hotspots, and female preference in the organization of lek mating systems. *American Naturalist, 131*, 203–219.

Benedict, L., & Najar, N. A. (2019). Are commonly used metrics of bird song complexity concordant? *Auk, 136*, 1–11.

Berglund, A., Bisazza, A., & Pilastro, A. (1996). Armaments and ornaments: an evolutionary explanation of traits of dual utility. *Biological Journal of the Linnean Society, 58*, 385–399. https://doi.org/10.1006/bijl.1996.0043

Bostwick, K. S., & Prum, R. O. (2003). High-speed video analysis of wing-snapping in two manakin clades (Pipridae: Aves). *Journal of Experimental Biology, 206*, 3693–3706.

Botero, C. A., Rossman, R. J., Caro, L. M., Stenzler, L. M., Lovette, I. J., de Kort, S. R., & Vehrencamp, S. L. (2009). Syllable type consistency is related to age, social status and reproductive success in the tropical mockingbird. *Animal Behaviour, 77*, 701–706.

Bradbury, J. W. (1981). The evolution of leks. In R. D. Alexander & D. W. Tinkle (Eds.), *Natural selection and social behaviour: recent research and new theory* (pp. 138–169). Chiron Press.

Bradbury, J. W., & Vehrencamp, S. L. (2014). Complexity and behavioural ecology. *Behavioural Ecology, 25*, 435–442.

Byers, B. E., & Kroodsma, D. E. (2009). Female mate choice and songbird song repertoires. *Animal Behaviour, 77*, 13–22.

Cardoso, G. C., & Hu, Y. (2011). Birdsong performance and the evolution of simple (rather than elaborate) sexual signals. *American Naturalist, 178*, 679–686.

Costa, M., Goldberger, A.L., & Peng, C.-K. (2005). Multiscale entropy analysis of biological signals. *Physical Review E, 71*, 13–22.

Couzin, I. D., Krause, J., Franks, N. R., & Levin, S. A. (2005). Effective leadership and decision-making in animal groups on the move. *Nature, 433*, 513–516.

Csardi G., & Nepusz T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems 1695*.

Da Silva, M. L., Piqueria, 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*, 57–64.

Deneubourg, J. L., Aron, S., Goss, S., & Pasteels, J. M. (1990). The self-organizing exploratory pattern of the Argentine ant. *Journal of Insect Behaviour, 3*, 159–168.

Emlen, S. T., & Oring, L. W. (1977). Ecology, sexual selection, and the evolution of mating systems. *Science, 197*, 215–223.

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

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*, 68–81.

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology & Evolution, 7*, 1325–1330.

Fusani, L., Giordano, M., Day, L. B., & Schlinger, B. A. (2007). High-speed video analysis reveals individual variability in the courtship displays of male Golden-Collared Manakins. *Ethology, 113*, 964–972.

Gibson., R.M. (1996). Female choice in sage grouse: The roles of attracting and active comparison. *Behavioural Ecology and Sociobiology, 39*, 55–59.

Gibson, R. M., & Bradbury, J. W. (1985). Sexual selection in lekking sage grouse: phenotypic correlates of male mating success. *Behavioural Ecology & Sociobiology, 18*, 117–123.

Hewitt, S. E., Macdonald, D. W., & Dugdale, H.L. (2009). Context-dependent linear dominance hierarchies in social groups of European badgers, *Meles meles*. *Animal Behaviour, 77*, 161–169.

Hobson, E. A., John, D. J., Mcintosh, T. L., Avery, M.L. & Wright, T. F. (2015). The effect of social context and social scale on the perception of relationships in monk parakeets. *Current Zoology, 61*, 55–69.

Howe, H. F. (1979). Fear and frugivory. *American Naturalist, 114*, 925–31.

Huffman, D. A. (1952). A method for the construction of minimum-redundancy codes. *Proceedings of the I.R.E, 40*, 1098-1101.

Huxley, J. S. (1914). The courtship habits of the great crested grebe (*Podiceps cristatus*); with an addition to the theory of sexual selection. *Proceedings of the Zoological Society of London, 84*, 491–562.

Janisch, J., Perinot, E., & Fusani, L. Behavioural flexibility in the courtship dance of Golden-Collared Manakins, *Manacus vitellinus*. *Animal Behaviour 166*, 61–71.

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*, 414–420.

Kirkpatrick, M. (1982). Sexual selection and the evolution of female choice. *Evolution, 36*, 1–12.

Kodric-Brown, A., & Nicoletto, P. F. (2001). Female choice in the guppy (*Poecilia reticulata*): the interaction between male colour and display. *Behavioural Ecology & Sociobiology, 50*, 346–351.

Levey, D. J. (1987). Seed size and fruit-handling techniques of avian frugivores. *American Naturalist, 129,* 471–485.

Lill, A. (1974). Sexual behaviour of the lek-forming White-bearded Manakin (*Manacus manacus trinitatis* Hartert). *Ethology, 36,* 1–36.

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*, 166–172.

McDonald, D. B. (1989). Correlates of male mating success in a lekking bird with male-male cooperation. *Animal Behaviour, 37*, 1007–1022.

McDonald, D. B. (2007). Predicting fate from early connectivity in a social network. *Proceedings of the National Academy of Sciences, 104*, 10910–10914.

McDonald, D. B. (2010). A spatial dance to the music of time in the leks of Long-Tailed Manakins. *Advances in the Study of Behaviour 42*, 55–81.

McDonald, D. B., & Hobson, E. A. (2018). Edge weight variance: population genetic metrics for social network analysis. *Animal Behaviour, 136*, 239–250.

McDonald, D. B., & Potts, W. K. (1994). Cooperative display and relatedness among males in a lek-mating bird. *Science, 266*, 1030–1032.

McDonald, D. B., & Shizuka, D. (2013). Comparative transitive and temporal orderliness in dominance networks. *Behavioural Ecology 24*, 511-520.

Miles, M. C., & Fuxjager, M. J. (2018). Synergistic selection regimens drive the evolution of display complexity in birds of paradise. *Journal of Animal Ecology, 87*, 1149–1159.

Miller, E.H. (1988). Description of bird behaviour for comparative purposes. *Current Ornithology, 5*, 347–394.

Morales-Betancourt, J.A., & Castaño-Villa, G.J. (2018). Males in seemingly female-like plumage do not mimic females. *Journal of Avian Biology, e01467*.

Muller, Z., Cantor, M., I.C. Cuthill, I. C., & Harris, S. (2018). Giraffe social preferences are context dependent. *Animal Behaviour 146*, 37–49.

Nelson, X. J., Wilson, D. R., & Evans, C. S. (2008). Behavioural syndromes in stable social groups: an artifact of external constraints? *Ethology*, 114, 1154–1165.

Nooker, J. K., & Sandercock, B. K. (2008). Phenotypic correlates and survival consequences of male mating success in lek-mating greater prairie-chickens (*Tympanuchus cupido*). *Behavioural Ecology & Sociobiology, 62*, 1377–1388.

Ooms, J., & Google, Inc. (2018). brotli: A compression format optimized for the web. R package version 1.2. 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., & Hebets, E. A. (2016). New dimensions in animal communication: the case for complexity. *Current Opinion in Behavioural Sciences, 12*, 80–89.

Patricelli, G. L., Coleman, S. W. & Borgia, G. (2006). Male Satin Bowerbirds, *Ptilonorhynchus violaceus*, adjust their display intensity in response to female startling: an experiment with robotic females. *Animal Behaviour* 71, 49–59.

Pinter-Wollman, N., Hobson, E. A., Smith, J. E., Edelman, A. J., Shizuka, D., de Silva, S., Waters, J. S., Prager, S. D., Sasaki, T., Wittemyer, G., Fewell, J., & McDonald, D. B. (2014). The dynamics of animal social networks: analytical, conceptual, and theoretical advances. *Behavioural Ecology, 25,* 242–255.

Potts, W. K. (1984). The chorus-line hypothesis of maneuver coordination in avian flocks. *Nature, 309*, 344.

Prum, R., & Johnson, A. (1987). Display behaviour, foraging ecology, and systematics of the Golden-Winged Manakin (*Masius chrysopterus*). *Wilson Bulletin, 99*, 521–539.

R Core Team, (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reding, L., & Cummings, M. E. (2017). Context-dependent preferences vary by multicomponent signals in a swordtail. *Animal Behaviour, 129*, 237–247.

Reynolds, J. D., & Gross, M. R. (1990). Costs and benefits of female mate choice: is there a lek paradox? *American Naturalist, 136*, 230–243.

Rodríguez, R. L., Haen, C., Cocroft, R. B. & Fowler-Finn, K. D. 2012. Males adjust signaling effort based on female mate-preference cues. *Behavioural Ecology* 23, 1218–1225.

Ryder, T. B., & Durães, R. (2005). It’s not easy being green: using molt and morphological criteria to age and sex green-plumage manakins (Aves: Pipridae). *Ornitologia Neotropical, 16*, 481–491.

Ryder, T. B, McDonald, D. B., Blake, J. G., Parker, P. G., & Loiselle, B. A. (2008). Social networks in the lek-mating Wire-tailed Manakin (*Pipra filicauda*). *Proceedings of the Royal Society of London Series B: Biological Sciences, 275*, 1367–1374.

Sibly, R. M., Nott, H. M. R., & Fletcher, D. J. (1990). Splitting behaviour into bouts. *Animal Behaviour 39*, 63-69.

Sih, A., Bell, A, & and Johnson, J. C. (2004). Behavioural syndromes: an ecological and evolutionary overview. *Trends in Ecology & Evolution 19*, 372–378.

Snow, D. W., & Snow, B. K. (1992). Display of the Golden-winged Manakin *Masius* *chrysopterus*. *Bulletin of the British Ornithologists’ Club, 112*, 264–270.

Sonderegger, D. (2020). SiZer: Significant zero crossings. R package version 1.1-7. https://CRAN.R-project.org/package=SiZer

Tinbergen, N. (1963). On aims and methods of Ethology. *Ethology, 20*, 410–433.

Trainer, J. M., & McDonald, D. B. (1993). Vocal repertoire of the Long-Tailed Manakin and its relation to male-male cooperation. *Condor, 95,* 769–81.

Uriarte, M., Anciães, M. da Silva, M. T. B., Rubim, P., E. Johnson, E., & Bruna, E. M. (2011). Disentangling the drivers of reduced long-distance seed dispersal by birds in an experimentally fragmented landscape. *Ecology 92*, 924–937.

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

Vanderbilt, C. C., J.P. 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*, 213–19.

Wheelwright, N. T. (1991). How long do fruit-eating birds stay in the plants where they feed? *Biotropica 23*, 29–40.

Ziv, J., & Lempel, A. (1977). A universal algorithm for sequential data compression. *IEEE Transactions of Information Theory, 12*, 337–343.

Zuk, M., Thornhill, R., Ligon, J. D., Johnson, K., Austad, S., Ligon, S. H., Thornhill, N. W., & Costin, C. (1990). The role of male ornaments and courtship behaviour in female mate choice of red jungle fowl. *American Naturalist, 136*, 459–473.

**ACKNOWLEDGMENTS**

The Mindo Cloudforest Foundation (MCF) allowed us to use their Milpe Sanctuary for our study. Brian Krohnke and Maria José Guerrero of MCF were extremely helpful in providing logistical support. Support from the Fulbright Foundation, and the University of Wyoming's Flittie sabbatical fund helped enhance the field efforts. We are grateful to Grady Harris, Will McDonald, and many field assistants for hard work in the field that provided the scaffold on which we built our research. The University of Wyoming Education Abroad Program provided some of the travel funds for the authors. Gracias tambien a la vida silvestre, al pueblo y al gobierno del Ecuador.

**Compliance with Ethical Standards**

*Conflict of Interest*

The authors declare that they have no conflict of interest.

*Ethical Approval*

All applicable international, national, and institutional guidelines for the care and use of animals were followed. Capture by mist nets, handling and banding were approved by the University of Wyoming's Institutional Animal Use Committee (IACUC protocol #20160602DM00242-01) for each of the years of fieldwork.

**Supplementary Materials**

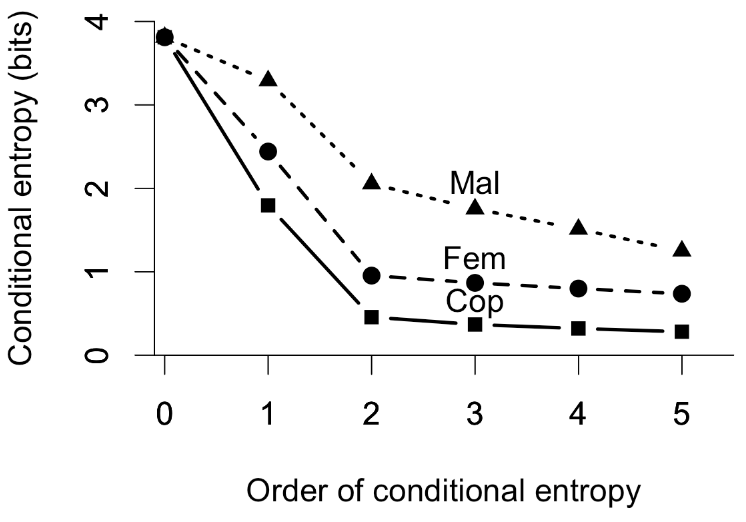


Fig. S1. Higher order (conditional) entropies for Mal, Fem, and Cop bouts. Order 1 is the first-order entropy of Fig. 2. Note that Mal continues to drop for higher-order entropies beyond Order 2. That difference suggests that in Fem and Cop bouts, the current element depends largely on the preceding element, whereas in Mal bouts, elements depend, in complex ways, on earlier elements. Nevertheless, the sharpest drop for all three is from Order 1 to Order 2, suggesting that the sequences are a reasonable fit to a first-order Markov chain process.

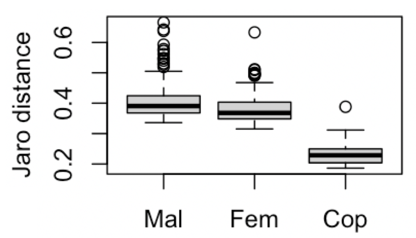


Fig. S2. Boxplots for Jaro distances in 198 Mal bouts, 100 Fem bouts and 14 Cop bouts. Cop bouts are considerably more similar to each other (low Jaro distances), than are either Mal or Fem bouts.

Table S1. List of 41 unmodified BORIS behavioural elements that were abbreviated, omitted or combined (indented in italics) to create the 21 numbered variables (ethogram elements) listed in Table 1.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Behaviour value in BORIS | Frequency | Abbreviation for analyses |
| 1 | Attempted Copulation | 11 | 1. AttCop |
| 2 | Bird2 ALAD | 55 | 2. B2AL |
| 3 | Bird2 BowLeft | 695 | 3. B2Bw |
| 4 | Bird2 BowRight | 677 | B2Bw |
| 5 | Bird2 Half-bow Left | 26 |  |
| 6 | Bird2 Half-bow Right | 21 |  |
| 7 | Bird2 HeadDownBowing | 56 |  |
| 8 | Bird2 Mixed Element | 13 |  |
| 9 | Bird2 NeckTwist | 1693 | 4. B2Nk |
| 10 | Bird2 On-log NO display | 79 |  |
| 11 | Bird2 SLAD | 1 |  |
| 12 | Bird2 TafLF\_Off | 28 |  |
| 13 | Bird2 TafLF\_On | 26 |  |
| 14 | Copulation | 24 | 5. Cop |
| 15 | End | 486 | 6. End |
| 16 | Female Looking Away | 1736 |  |
| 17 | Female Movement | 978 |  |
| 18 | Female Off Log | 320 | 7. Fff |
| 19 | Female On Log | 320 | 8. FOn |
| 20 | Female ResponseToALAD | 188 |  |
| 21 | Female Tracking Male | 1924 |  |
| 22 | FemaleSwitch | 406 |  |
| 23 | Male1 ALAD | 1104 | 9. ALAD |
| 24 | Male1 BowLeft | 8088 | 10. Bow |
| 25 | Male1 BowRight | 7950 | Bow |
| 26 | Male1 Half-bow Left | 296 | 11. HafB |
| 27 | Male1 Half-bow Right | 258 | HafB |
| 28 | Male1 HeadDownBowing | 2361 | 12. HdBw |
| 29 | Male1 Metronome\_Left | 11 | 13. Metr |
| 30 | Male1 Metronome\_Right | 25 | Metr |
| 31 | Male1 Mixed Element | 361 | 14. Mix |
| 32 | Male1 NeckTwist | 5618 | 15. Neck |
| 33 | Male1 Off Log | 41 |  |
| 34 | Male1 On Log | 39 |  |
| 35 | Male1 On Log No Display | 2225 | 16. Zero |
| 36 | Male1 Other Behaviour | 2142 | 17. OthX |
| 37 | Male1 SLAD | 655 | 18. SLAD |
| 38 | Male1 Switch | 1145 | 19. Swtc |
| 39 | Male1 TafLF\_Off | 1193 | 20. Taf |
| 40 | Male1 TafLF\_On | 1223 | Taf |
| 41 | Start | 486 | 21. Start |

Notes for 2022-2023 revision.

Fig. 1. Four measures of the gradient of complexity from complex male-only courtship display bouts (Mal) to courtship displays with one or more females present but no copulation (Fem) to least complex in courtship displays ending in a copulation (Cop)  
a. Elem count/BoutLength ratio (output of 10-Oct-22). Ratio of (distinct element count)/(string length) here.   
  
 Chart, box and whisker chart

Description automatically generated

b. Entropy is as much a measure of surprise as of complexity. Entropy goes from low (few surprises) in display bouts ending in a copulation (Cop) to high in display bouts involving only males (Mal), with intermediate entropy for display bouts attended by one or more females that did not end in a copulation (Fem).   
  
 Chart, box and whisker chart

Description automatically generated

c. Compression string length (raw or ratio?). Ratio here. A courtship bout whose string of display elements can be greatly compressed (high compression ratio on the ordinate), such as the Cop display bouts, can be viewed as less complex than ones that are not very compressible, such as the Mal display bouts.   
  
 Chart, box and whisker chart

Description automatically generated  
  
 d. Algorithmic complexity, as an approximation of Kolmogorov complexity. Again, the   
 complexity gradient goes from low complexity for display bouts ending in a copulation   
 (Cop), to high complexity in display bouts involving only males (Mal).

Chart, box and whisker chart

Description automatically generated

From Liam's email of 6-Oct-22

"Focusing on the compressibility score as the central value and explicitly comparing it to simple measures like "repertoire size" might make for a very simple+impactful presentation of the data.

A central figure might be as simple as a four subplot bar plot:

(A) Repertoire sizes. Mal vs. Fem (exclude cop) vs. Cop.

(B) Entropy. Mal vs. Fem (exclude cop) vs. Cop

(C) Length of compression string. Mal vs. Fem (exclude cop) vs. Cop

(D) If any different than (C), Algorithmic complexity from acss. M vs. F vs. C

Then another figure (or worked into the above) parsing within vs. between individuals.

Plot (A) gives us the background value that most investigators might have looked for.

Plot (B) tells us something about predictability -- from the view of the female, how "surprised" she might possibly be, given the available male display elements.

Plot (C) tells us something about an approximation of complexity -- from the view of the male, how simple is the display to produce (aesthetic structure, repetition, etc.).

Plot (D) bolsters plot (C) with a slightly different estimator of complexity."

Questions/comments for Liam (22-Oct-22)

Should we call the 11 elements "events" rather than "elements"? Some (Cop, AttC, FOn, Fff, Mix, Othr, Zro) are not display elements per se (especially in the Schaedler comparative scheme).

Are the (almost impercepible, except in slomo) neck bobs and tail wags after an ALAD worth attomizing?

Vault (male vaults over female when both are on the log): does it occur in context other than ALAD?

Gradient from Neck twist (*Neck*) to Head-down Bow (*HdBw*)

Nick's lower canopy flybys?

"Kichup" sound in ALAD is a vocalization. See "1519\_Taf&kichup ALAD voc.mp4" & "1519\_CreekLog\_11\_18\_17\_02,mp4"

I need to check on "Gaping", Nick's fly-bys and wing-flicks. Check email of 22-Oct-22 and new voucher videos in Dropbox.

Hasenjager, Matthew J., Victoria R. Franks, and Ellouise Leadbeater. “From Dyads to Collectives: A Review of Honeybee Signalling.” Behavioral Ecology and Sociobiology 76, no. 9 (September 2022): 124. <https://doi.org/10.1007/s00265-022-03218-1> Zotero.

Bosshard, Alexandra B., Maël Leroux, Nicholas A. Lester, Balthasar Bickel, Sabine Stoll, and Simon W. Townsend. “From Collocations to Call-Ocations: Using Linguistic Methods to Quantify Animal Call Combinations.” Behavioral Ecology and Sociobiology 76, no. 9 (September 2022): 122. <https://doi.org/10.1007/s00265-022-03224-3>. Zotero