Social Network Dynamics of Budgerigars and their Influence on Vocal Learning

A Thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfillment of the requirements for the BS-MS Dual Degree Programme

by

Nakul Prashant Wewhare

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA.

Date: March, 2025

Under the guidance of

Supervisor: Dr Anand Krishnan,

Affiliation of Supervisor: Jawaharlal Nehru Centre For Advanced Scientific Research

From May 2024 to Mar 2025

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

Certificate

This is to certify that this dissertation entitled “Social network dynamics of Budgerigars and their influence on vocal learning towards” the partial fulfillment of the BS-MS dual degree program at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Nakul Prashant Wewhare at Jawaharlal Nehru Centre For Advanced Scientific Research under the supervision of Dr Anand Krishnan, Assistant Professor, Evolutionary and Organismal Biology Unit, during the academic year 2024-2025.

Dr Anand Krishnan
Assistant Professor,
Evolutionary and Organismal Biology Unit,
Jawaharlal Nehru Centre For Advanced Scientific Research

Committee:
Dr Anand Krishnan (Supervisor)
Dr Raghav Rajan (Expert)

Declaration

I hereby declare that the matter embodied in the report entitled “Social network dynamics of Budgerigars and their influence on vocal learning” are the results of the work carried out by me at the Evolutionary and Organismal Biology Unit, Jawaharlal Nehru Centre For Advanced Scientific Research (JNCASR), Banglore, under the supervision of Dr Anand Krishnan, and the same has not been submitted elsewhere for any other degree. Wherever others contribute, every effort is made to indicate this clearly, with due reference to the literature and acknowledgment of collaborative research and discussions.

Nakul Prashant Wewhare
Roll Number: 20201166
BS-MS
IISER Pune
Date: 14/03/2025

Table of Contents

List of Figures

List of Tables

Abstract

Social living involves a tradeoff between cooperation and competition, often giving rise to structured dominance hierarchies and intricate vocal communication. Budgerigars (*Melopsittacus undulatus*), with their capacity for open-ended vocal learning and complex social behaviors, offer a prime opportunity to examine these processes. In a mixed-sex captive flock studied over four weeks, we combined social network analysis of affiliative, aggressive, and display interactions with acoustic analyses of warble song syntax and contact call-like notes. Our results reveal a shallow yet consistent dominance hierarchy, where females generally outrank males. Notably, male displays typically used in courtship, also appear in male-male interactions as submission or appeasement signals, thereby mediating dominance relations. Warble song syntax and contact call repertoires exhibit strong individual signatures, yet both are shaped by social context. Males displaying to the same female showed convergence in their contact call repertoire, possibly signaling affiliation, while concurrently diverging in warble song syntax, thus maintaining distinct identities. Dominance status further influenced contact call structure, as more dominant males occupied central positions in the contact call similarity network. These findings suggest that subordinate males selectively converge with higher-ranked individuals, potentially as an affiliative signal. This study provides the first evidence for acoustic convergence as a mediator in dominance hierarchy, and a detailed account of their social organization in budgerigars. Overall, it underscores how intricate social networks and open-ended vocal learning intersect in parrots, informing broader discussions on the evolution of social communication systems.

.

Acknowledgments

Joining Dr. Anand Krishnan's lab in my second-year summer was probably the best decision of my life. I owe Anand most of what I have learned about how to do good science. While he has taught me to be quantitative and critical in my research, he has also emphasized having fun and staying curious in the process. The amount of support and effort he puts into mentoring each of his students is unparalleled by anyone I have heard of. His belief in my abilities has made me believe in myself and be more confident. I have called Anand's lab home for the last three years, and I am very grateful for having had the privilege to do so.

What makes Anand's lab the best are its people, and I am very thankful to all of them—Saro and Alakananda for always giving good life advice and support, and for their help in the initial setup of the bird facility and care of the birds; and Ahana for her help in the later stages. A big thanks to these three for all the fun, pranks, and parties we had over the last year, which are the memories I'll cherish most from my Master's thesis year. I am grateful to all the other members of the lab who have made my life happier in some way - Priya, Kezia, Amey, Abhinav, Arjun, Abhijith, Siddharth, and Sudhanva, and most importantly, to Muniraju Anna for taking care of the birds and sharing detailed histories of JNC.

I am very grateful to Dr. Kaja Wierichka and Dr. Ben Whittaker, as both have gone out of their way to mentor and help me, whether it was coding, statistics, scientific writing, or simply how to articulate and streamline my ideas.

I believe what makes IISER Pune a great place are its students. And for me, the people of 116 made IISER home. I am very thankful to all the people with whom I had any meaningful interaction. I am especially thankful to my roommate Anish, whose constant life adventures and good humor made my life far more entertaining than I couldn't have asked for a better roommate. He bore and tolerated all my quirks yet has always been a great friend to me. Akilan, for countless chats and discussions about everything eco-evo, which have made me understand and appreciate science more than any course at IISER ever did. Iyer, for all the random and fun chats about any topic under the sun. Vihang, for his sagely advice. Ujwal, for his camaraderie in all the cycling adventures big and small—and to my cycle, which kept me going whenever life felt like too much of an uphill climb.

Most importantly, Shashwati, whose support I could always rely on, believed in me more than I or anyone else ever did. She has made me a happier and more compassionate person than I ever was before. IISER would truly not have been the same without her. She has helped me so much with my projects that she could give a talk or a poster on them as well as I could. I have great respect for her personal and work ethic, which has spurred me to improve my own. Thanks for everything and for always being there for me.

Lastly, I am eternally grateful to my family Mummy, Papa, and Shubhi who have always supported me and never stopped me from doing anything, whether it was storing specimens in the common home freezer or going out at night to record dogs barking. While they have been encouraging of all my endeavors, they have also given me complete personal and academic freedom, which I greatly cherish and appreciate..

Contributions

Contributor name	Contributor role
Nakul Wewhare, Anand Krishnan	Conceptualization Ideas
Nakul Wewhare	Methodology
Nakul Wewhare	Software
Nakul Wewhare	Validation
Nakul Wewhare	Formal analysis
Nakul Wewhare	Investigation
Anand Krishnan	Resources
Nakul Wewhare	Data Curation
Nakul Wewhare	Writing - original draft preparation
Nakul Wewhare, Anand Krishnan	Writing - review and editing
Nakul Wewhare	Visualization
Anand Krishnan	Supervision
Nakul Wewhare, Anand Krishnan	Project administration
Anand Krishnan	Funding acquisition

Chapter 1: Introduction

Social or group living in animals confers numerous advantages, including collective vigilance against predators, shared information about resources, and greater opportunities to find mates. At the same time, group living amplifies competition, introducing a tradeoff between the benefits of cooperation versus conflict (Chase, 1974; Landau, 1951; Tibbetts et al., 2022). In stable groups with repeated interactions, hierarchies substantially reduce the costs and risks of ongoing aggression. Dominance hierarchies typically emerge when animals live in close proximity, leading to frequent encounters. For example, in chickens (*Gallus gallus domesticus*), these interactions form the basis for the classic “pecking order” (Guhl, 1956; Rushen, 1982). By establishing clear rank relationships, group members minimize costly and potentially injurious physical conflict over food or mates (Chase, 1974).

A key property of dominance hierarchies is transitivity: if A dominates B (by “winning” an encounter) and B dominates C, A generally dominates C as well. Winners reinforce their dominance over losers through repeated agonistic interactions, fights, chases, or displays. Once established, rank of each individual strongly influences its survival and reproductive success (Boucherie et al., 2022; Hobson et al., 2021; Shizuka & McDonald, 2012). Dominant individuals often secure better foraging spots, safer roost sites, more (or higher quality) mates, and exert more influence on group movements (Ang & Manica, 2010; Hobson, 2022). By contrast, subordinates may pay a cost in terms of delayed breeding opportunities, but still benefit from group membership via decreased predation risk or eventual access to resources (when the hierarchy shifts) (Hobson, 2022). After the initial phase of rank sorting, maintaining the hierarchy usually involves ritualized signals rather than escalated aggression. Dominant individuals may deliver brief reminders, threats, or chases to subordinates who overstep, whereas subordinates respond with appeasement gestures. Communication is thus pivotal in stabilizing hierarchies (Preuschoft et al., 2001, Tibbetts et al., 2022. Many species possess “badges of status”, such as unique color patterns, scents, or vocal cues, which enable rapid recognition of an

individual's social position. For example, in birds, conspicuous throat patches help to convey status (Dapporto et al., 2007; Møller, 1987). Chemical signals, for instance, convey dominance in wasps, whereas visual displays can assert hierarchy among gelada baboons or facilitate courtship in birds (Emory, 1976; Tibbetts et al., 2022). Vocal signals frequently play a major role, as they are flexible and can swiftly reflect a caller's identity, motivation, or rank even over relatively long distances. This combination of properties enables other group members to assess dominance relationships rapidly, thereby minimizing the need for costly physical contests (Preuschoft et al., 2001, Tibbetts et al., 2022).

1.1 Vocal communication for cohesion

Stable social groups often rely on vocal signals to maintain proximity and respond to changes in the environment. Bottlenose dolphins *(Tursiops truncatus)* use individually unique "signature whistles" to recognize one another over large distances, thereby preserving social bonds even when group members are apart (King et al., 2014). Colonial birds such as penguins *(Aptenodytes forsteri)* employ individual-specific calls so that parents and chicks can reunite in crowded rookeries (Aubin & Jouventin, 1998). African Elephants (*Loxodonta africana*) emit low-frequency rumbles to coordinate group departures, and others rumble back in a form of collective decision-making (O'Connell-Rodwell et al., 2024). Meerkats *(Suricata suricatta)* and primates in dense habitats emit contact calls to confirm each other's presence, ensuring that group members remain within communication range and reducing the likelihood of accidental separation (Demartsev et al., 2018).

1.2 Using vocalizations to assert rank

Beyond cohesion, vocal signals play a pivotal role in forming and maintaining dominance hierarchies. They may function as aggressive, submissive or appeasement cues that reduce the frequency of escalated conflict. In chimpanzees *(Pan troglodytes)*, "pant-grunts" acknowledge a higher-ranking individual and help avert further aggression, whereas baboons (*Papio cynocephalus ursinus*) use low grunts from dominant females to reassure subordinates and pave the way for post-conflict grooming (Cheney et al., 1995; Van Lawick-Goodall, 1968). The "giggle"

of the spotted hyena *(Crocuta crocuta)* encodes individual identity, rank, and emotional state, enabling listeners to determine whether they should approach or avoid the caller (Mathevon et al., 2010). Similarly, in crowded roosts, Egyptian fruit bats *(Rousettus aegyptiacus)* use specific social calls to argue over resources rather than constantly engage in physical fighting (Prat et al., 2016). Among birds, the "pecking order" among chickens *(Gallus gallus domesticus)* relies in part on threat cackles and submissive squawks that limit repeated brawls (Marino, 2017), whereas large-billed crows *(Corvus macrorhynchos)* use long repetitive sequences of calls to signal dominance (Kondo & Hiraiwa-Hasegawa, 2015), and songbirds overlap or match each other's songs in territorial contests to signal status and de-escalate conflict (Hyman, 2003). Across these examples, vocalizations serve to clarify the hierarchy, reduce risk of injuries, and enable rapid signals of dominance and submission.

1.3 Vocal convergence or vocal accommodation in social animals

Beyond dominance and conflict resolution, vocalizations also play a crucial role in recognition of group members. In many species that exhibit vocal plasticity, group living promotes the convergence of call structures to enable recognition of group identity, aiding in the identification of group versus non-group members (Ruch et al., 2018). Convergent "dialects" have been documented in songbirds such as finches and sparrows, where flock-mates adopt similar song signatures (Goodwin & Podos, 2014; Podos & Warren, 2007). Similarly, diverse parrots and canaries *(Serinus canaria domestica)* alter their vocalizations in parallel with increasing social affiliation, a pattern that extends to mammals as well (Tyack, 2020). Greater spear-nosed bats *(Phyllostomus hastatus)*, for example, form female-dominated groups in which each group maintains a distinct contact-call dialect and preferentially responds to matching calls (Boughman, 1998). Chimpanzees *(Pan troglodytes)* from the same community exhibit convergence in "pant-hoots," and newly integrated individuals gradually shift toward the local variant (Mitani & Gros-Louis, 1998). Marmoset (*Callithrix jacchus*) pairs likewise converge in pitch and trill patterns once bonded (Zürcher et al., 2021). Among marine mammals, killer whales *(Orcinus orca)* use pod-specific call repertoires transmitted matrilineally (Miller & Bain, 2000), whereas sperm whales *(Physeter macrocephalus)* rely on clan-specific click

sequences (“codas”) for group cohesion (Hersh et al., 2022). Bottlenose dolphins *(Tursiops truncatus)* primarily develop stable signature whistles during early social bonding, but minor acoustic adjustments can occur within alliances (Janik et al., 2006). These convergent signals serve multiple functions: they reinforce collective identity, reduce conflict, and may also potentially reinforce dominance relationships (Ruch et al., 2018; Wright & Dahlin, 2018). In principle, subordinates could align their calls with those of dominant individuals to signal affiliation, thereby negotiating social rank through shifts in vocal behavior rather than relying on relatively costly displays of aggression or submission. By emphasizing shared acoustic features, socially cohesive groups benefit from improved coordination and recognition.

1.4 Linking vocal complexity and social structure

As group size grows and social networks become more intricate, the benefits and occurrences of diverse, complex vocal signals increase (the “social complexity hypothesis”) (Bouchet et al., 2013; Krams et al., 2012; Peckre et al., 2019). Species such as parrots, dolphins, and songbirds, all open-ended vocal learners (which learn throughout their lifetime), may continuously revise their calls in response to evolving social bonds, changes in group composition, or shifts in dominance. Vocal flexibility enables open-ended learners to manage multiple communicative tasks simultaneously, from individual recognition and group membership to status signaling (Tyack, 2020). Nevertheless, many questions remain about how animals balance these different functions, particularly when vocal signals can be modified well into adulthood. It remains unclear whether ongoing changes in flock composition or rank could reshape an individual’s repertoire. Investigating the links between social environment and vocal signaling may help illuminate how group-living animals navigate the complexities of cooperation and competition through learned acoustic signals (Preuschoft et al., 2001; Smeele et al., 2024).

1.5 Parrots as a model system for social vocal learning

Parrots (order Psittaciformes) exhibit a combination of sophisticated social organization and open-ended vocal learning, making them valuable models to study the evolution and function of complex social communication. Like humans, but unlike many better-studied songbird systems, parrots acquire new vocalizations beyond early development, constantly modifying and adding to their repertoires throughout adulthood (Bradbury & Balsby, 2016; Wright & Dahlin, 2018). Many species form large, dynamic flocks that exhibit “fission-fusion” patterns (Bradbury & Balsby, 2016; Hobson et al., 2014; Vehrencamp et al., 2003). However, smaller subgroups within these flocks may remain stable, thereby facilitating prolonged social interactions and structured dominance hierarchies (Penndorf et al., 2023, 2024; Vehrencamp et al., 2003). This dual feature of highly flexible vocal learning and complex flock organization is rare in the animal world (Bradbury & Balsby, 2016).

Individually distinctive “contact calls” are a well-studied hallmark of parrot communication. Flock-mates often learn or imitate one another’s calls, likely supporting rapid individual recognition as flocks split or merge (Buhrman-Deever et al., 2008; Dahlin et al., 2014; Scarl & Bradbury, 2009; Wright & Dahlin, 2018). Some parrots also produce more elaborate vocal sequences variously termed “warbles” or “songs” that combine diverse note types, and may encode context-specific information such as group membership or affiliative bonds (Madabhushi, Wewhare, et al., 2023). This acoustic complexity is further enhanced by parrot cognitive abilities, including advanced memory and problem-solving skills (Auersperg & Von Bayern, 2019). Collectively, the attributes of open-ended learning, social complexity, and strong cognitive capacities indicate that parrots are prime subjects to examine how complex vocal signals develop, adapt, and mediate social relationships.

Field studies of wild parrots, however, face serious challenges stemming from the birds’ wide-ranging daily movements and high mobility (Dahlin et al., 2024; Hobson et al., 2014). Many parrot species forage in tall tree canopies or traverse extensive savannas in search of spatially patchy food sources, making it difficult to locate, tag, and resight individuals consistently (Brockway, 1964a; Hobson et al., 2014). Nomadic species such as budgerigars (*Melopsittacus undulatus*) move long

distances, changing roost sites frequently and thwarting efforts to track social interactions over time (Wyndham, 1983). When multiple individuals within a flock vocalize simultaneously, it becomes very difficult to record acoustic signals of specific birds or even to confirm caller identity, especially if individuals are unmarked or move quickly out of view (Bradbury & Balsby, 2016; Hobson et al., 2014). Only a handful of field projects such as those on urban Sulphur-crested cockatoos (*Cacatua galerita*) have successfully tracked marked birds for sustained intervals, and even then, acquiring high-quality vocal and behavioral data in tandem remains an obstacle (Penndorf et al., 2024).

Captive studies alleviate many of these constraints, and have therefore been instrumental in studies of complexity and dynamics in parrot vocal learning (Dahlin et al., 2024; Dooling et al., 2006; Farabaugh et al., 1994; Madabhushi, Wewhare, et al., 2023). In captive populations, researchers could systematically record known individuals, monitor group interactions continuously, and manipulate social contexts by introducing or removing specific birds. This approach has clarified, for instance, how parrots develop “vocal passwords” within stable flocks, or how readily they adjust call structure to include newly arrived group members (Dahlin et al., 2014). Dominance hierarchies, courtship, and other social processes are more analytically and experimentally tractable, and investigators may examine the often-opposing effects of social status, vocal signaling, and group cohesion under relatively controlled conditions (Hobson & DeDeo, 2015). Although captive populations lack certain natural pressures (such as predation risk) and do not undertake large-scale nomadic movements, these controlled experiments provide crucial foundations to understanding the mechanisms of parrot communication (Dahlin et al., 2014; Farabaugh et al., 1994; Manabe et al., 2008; Moravec et al., 2006).

By integrating insights from both captive and free-ranging populations, we begin to understand how parrots use and adapt their vocal repertoires in dynamic social environments, and how these processes compare with more traditional avian models of vocal learning (e.g., oscine songbirds). Parrots possess intricate coordination of calls in flocks (Bradbury & Balsby, 2016; Wright & Dahlin, 2018), the modification of individual repertoires during mating and pair-bonding (Hile et al., 2000; Wewhare & Krishnan, 2024), and the potential to communicate with specific individuals in a flock

through short-term vocal convergence (Balsby et al., 2012; Balsby & Scarl, 2008). This highlights their power as a study system in animal communication and social learning, offering a valuable parallel to human speech development. In both systems, adult and juvenile learning, flexible usage of signals, and strong social influences shape the emergence and maintenance of complex vocal cultures (Bradbury & Balsby, 2016; Wright & Dahlin, 2018).

1.6 The Budgerigar

The budgerigar *(Melopsittacus undulatus)*, a small Australian psittaciform, is one of the best-studied parrots due to its social nature, open-ended vocal learning, and ease of breeding under captive conditions (Wyndham, 1980). In the wild, budgerigars roam the Australian interior in large, often nomadic flocks; in captivity, they maintain many of their natural behaviors, readily forming mixed-sex groups that facilitate experimental studies of how social organization shapes vocal learning (Brockway, 1964a; Wyndham, 1980, 1983). This combination of sociability and adaptability makes budgerigars exceptionally suitable to examine the relationship between social structure and learning of complex acoustic repertoires (Brockway, 1964a, 1964b).

Budgerigars produce two principal vocal types. Contact calls are short, loud, and frequency-modulated signals that both sexes use to maintain group cohesion and recognize individuals. These calls encode individual and group identity within their characteristic modulation patterns and often converge within a flock through active learning (Bartlett & Slater, 1999; Farabaugh et al., 1992, 1994). An individual budgerigar may produce several subtypes of contact calls, each with a distinct frequency modulation contour, some of which may be shared across a subgroup or used only by certain individuals. Even within a particular subtype, individual identity remains traceable even when the broader group converges (Tu et al., 2011; Zhao et al., 2023). By contrast, warble song is a longer, more elaborate vocal sequence, emitted primarily by males during courtship or during social arousal (Brockway, 1964a; Farabaugh et al., 1992). Early observations noted that warble sequences integrate various note types, including short “contact call-like” notes that resemble the bird’s standalone contact calls, but are emitted at lower amplitude and with

slightly altered durations (Farabaugh et al., 1992; Tu et al., 2011). Subsequent studies confirmed that warble syntax also converges among cohabiting birds, indicating that social factors influence both discrete call elements and the arrangement of notes into sequences (Madabhushi et al., 2023).

Historically, the importance of warble song in mate attraction and pair bonding was established during the 1960s and 1970s (Brockway, 1967), with research showing that male warbling can stimulate ovarian maturation in females (Brockway, 1965). Laboratory and aviary experiments in the 1990s further demonstrated that when unfamiliar budgerigars are housed together, they develop shared contact calls (Farabaugh et al., 1994; Hile et al., 2000). The following decade saw a shift toward pair-bonding studies, revealing that male budgerigars often imitate their female partner's contact call, and that females preferentially mate with males whose calls closely match their own (Moravec et al., 2006). In the 2010s, work by Christine Dahlin and others tested the "password" hypothesis, concluding that shared calls reflect affiliative processes, rather than a recognition mechanism to exclude outside individuals from a flock (Dahlin et al., 2014). More recently, Madabhushi et al. (2023) found evidence of vocal convergence in higher-order structure (syntax) of the warble song, resulting in a merged, colony-wide repertoire.

Despite these findings, many questions remain about how courtship, dominance, and group identity interact at the level of larger, mixed-sex flocks. Past studies often focused on a single dimension at a time, e.g. male-female vocal matching, or vocal behavior in female-only flocks (Hile et al., 2000; Hile & Striedter, 2000). However, free-ranging budgerigars likely manage multiple social interactions simultaneously. It is unknown, for instance, whether dominant individuals lead or shape the group's call repertoire, or whether subordinates adjust warble syntax to either overlap with or avoid competing with more dominant birds. In this case, similar social interactions could have very different consequences for vocal learning. Moreover, it remains unclear whether the patterns of convergence in both contact calls and higher-order warble structure are modulated by social rank, pair-bonding status, or male-male affiliative interactions. Most studies to date have treated call structure and warble syntax as separate phenomena, even though they may function as coordinated parts of a broader communication system. Understanding how these different levels of

vocal signals co-vary or diverge under the same social conditions would clarify whether budgerigars use similar social cues to shape both call usage and song syntax, or if each domain follows distinct rules and functions.

Investigation of mixed-sex flocks where budgerigars simultaneously engage in courtship, competition, and group interactions offers a promising path for uncovering links between vocal plasticity and social organization. This work will advance our understanding of how social complexity and vocal complexity co-evolve in parrots, and, more broadly, will illuminate fundamental processes that underlie sophisticated communication systems in other open-ended vocal learners.

1.7 Present study: aims and hypotheses

Building on prior research into budgerigar social organization and vocal flexibility, we conducted a four-week investigation of a mixed-sex captive flock to examine how hierarchical social structures intersect with vocal learning and communication. Our first objective was to characterize the dominance hierarchy of budgerigars. Studies of other parrots indicate that dominance may be moderately steep and also fluid, especially in species that exhibit fission-fusion dynamics. In Sulphur-crested cockatoos *(Cacatua galerita)*, dominance rank appears partly segregated by sex (male > female) (Penndorf et al., 2024), whereas monk parakeets *(Myiopsitta monachus)* do not show strong sex-based partitioning (Hobson et al., 2014). We, therefore, tested whether male and female budgerigars exhibit separate or overlapping hierarchies and whether sex predicts differences in rank stability or aggression rates.

To our knowledge, there is no detailed analysis of steepness and linearity in budgerigar dominance hierarchies. Assessing these properties may offer insights into whether a hierarchy is strongly linear with clear, consistent rank orders or if it is more flexible, permitting faster rank shifts when social or environmental conditions change. In addition, although male budgerigars are known to perform courtship displays toward females, such displays also frequently occur in male-male interactions (Abbassi & Burley, 2012). We explored the possibility that these displays in same-sex contexts serve functions beyond reproductive signaling; for instance, as

submissive or appeasement gestures within the dominance hierarchy. Understanding whether males adapt courtship-like displays to navigate male-male relationships addresses a key gap in our knowledge of how budgerigars manage multifaceted social demands through vocal and behavioral signals.

In the second part of this study, we investigated vocal behavior on two levels: (1) warble song syntax produced by males, and (2) the spectral structure and relative use of contact call-like elements (“b notes” henceforth) embedded within the warble. Prior research suggests budgerigars can learn or modify both call structure and sequence syntax in response to changing social contexts (Dahlin et al., 2014; Farabaugh et al., 1994; Madabhushi, Wewhare, et al., 2023). We therefore predicted that individual budgerigars would display distinctive “signatures” in both warble syntax and b notes, while also showing some degree of temporal drift or convergence over the study period. By comparing vocal distance metrics to social interactions, we examined whether metrics derived from social structure like the frequency of direct male-male displays or male displays to the same females correlated with vocal similarity metrics. Additionally, we aimed to see whether highly ranked birds occupied a central “node” in the acoustic similarity network, which could be interpreted as subordinates converging on vocalization of dominants in a rank-dependent manner.

To address these questions, we (a) video-recorded a captive colony daily for four weeks to quantify social behaviors (affiliative, aggressive, display, and copulatory) and to construct weekly social networks, (b) analyzed male warble songs in a soundproof enclosure to measure syntax metrics, and (c) extracted b notes from the same data to assess contact call similarity through dynamic time warping and clustering. By integrating social network analysis and detailed acoustic data, we tested whether dominance rank is correlated to patterns of vocal convergence, whether sex differences shape the hierarchy and vocal usage, and whether display interactions among males (whether direct or via displays to the same females) predict the degree of acoustic overlap. Through these approaches, our study provides comprehensive insight into how budgerigars negotiate both social hierarchy and vocal identity in a complex captive setting, shedding light on the intersection between group living and social learning more broadly.

Chapter 2: Methods

2.1 Study Animals

We carried out our experiments on a colony of captive budgerigars (*Melopsittacus undulatus)* held at the Jawaharlal Nehru Centre for Advanced Scientific Research (JNCASR) in Bangalore. All procedures were approved by the Institutional Animal Ethics Committee (protocol number AK 001) in accordance with the guidelines of the Committee for the Purpose of Control and Supervision of Experiments on Animals (CPCSEA, New Delhi).

The colony consisted of 13 birds (6 females, 7 males) purchased from a local breeder. Birds were initially sexed using cere color, and we also confirmed this using genetic sexing (DNA extracted from their feathers). Each bird was ringed on both legs with metal rings bearing unique color combinations and printed numbers for identification. The birds were quarantined for one month and then acclimatized for another month before data collection. Data was collected over four weeks, from 30 October - 26 November 2024. The birds were housed together in a single colony cage measuring 6 × 6 × 6 ft, maintained at 25°C using an air conditioner and run on a 12 h:12 h light-dark cycle using fluorescent lamps and a timer switch. Water, commercial bird seed, and cuttlebone were provided ad libitum, and we provided boiled eggs, baby corn, chickpeas, coriander, and spinach leaves every few days as dietary supplements. For enrichment, we furnished the cage with a mixture of wooden and cotton-rope perches at various heights, along with multiple parrot-safe toys, replaced every few weeks to prevent habituation and introduce novelty.

2.2 Experimental setup

First, we filmed the entire colony for two hours daily (09:00–11:00) using a Sony PXW-Z90V 4K HDR XDCAM camcorder (Sony Corporation Minato, Tokyo, Japan), recording in XAVC QFHD format at 2160/25p and 25 fps. We chose this interval as peak activity typically began about one hour after the lights switched on at 08:00. The zoom and focus of the camera were positioned so that the 6 × 6 × 6 ft cage filled the frame, making the leg bands clearly visible for identification. This enabled us to monitor and quantify social interactions comprehensively. However, audio collected in the colony environment was extremely noisy, so we needed cleaner recordings to quantify individual repertoires and syntax. For this, we paired each male with a female inside a 75 × 75 × 75 cm acoustically isolated enclosure (Newtech Engineering Systems, Bengaluru) and recorded songs using an MKH 416-P48U3 directional microphone (Sennheiser Inc., Wedemark, Gernamy) positioned inside the sound enclosure. The microphone was connected to the camcorder, now placed outside the enclosure looking in through a glass observation window and recording synchronized video of the birds in the sound enclosure. Recording settings were switched to AVCHD at 1080/50i, with an audio sampling rate of 44.5 kHz. For subsequent acoustic analyses, we extracted the audio track from the video using VLC Media Player (VideoLAN, Paris, France) and converted it to WAV format for further analysis. Each male was paired with the female to whom he displayed most frequently (this was assessed before the study by observation of the colony during the acclimatization period). We kept these pairs consistent across the 4 weeks wherever possible, with some exceptions. Two males (SG and HE) shared the same female (WL), and male SB's partner was changed from female DG to female DB after DG died at the end of Week 3 (SB subsequently did not sing to the new female). One male, WT, never vocalized in the enclosure despite pairing him with several different female partners, though he did sing in the colony cage. The final audio data thus consisted of songs from six males over four weeks, except SB in Week 4. Pairs were recorded from 12:00 to 17:00, occasionally extended to 19:30 on the occasions where the male had not produced sufficient vocalizations. Each pair was recorded on the same weekday across the 4-week period, providing a consistent 1-week interval between recording sessions.

2.3 Analysis

2.3.1 Social structure

2.3.1.1 Video scoring

Although we collected daily video data, we analyzed data from the first 4 days of each week (our analyses suggested this was a sufficient data bin to address our primary questions). Across 4 weeks, this yielded 16 total days of behavioral observations for constructing social networks. Social networks can be generated either from patterns of spatial association or from direct interactions; however, proximity may not accurately reflect social choice in a confined environment such as an aviary. Therefore, we constructed our networks by focusing on observed *behavioral interactions*. From each 2-hour video for each day, we scored six 10-min segments (every other 10-min interval), resulting in 1 hour of analyzed video per day. Annotations were made using BORIS (version 8.27.7 2024-08-23) (Friard & Gamba, 2016) with an all-occurrence sampling strategy (Altmann, 1974). This was possible because all birds were always in view, and we frequently rewound the footage to ensure that every interaction was recorded. Every interaction was marked as a *point event*. If the same behavior between the same directional dyad (eg: aggression from male A to male B) recurred after 30 s, it was treated as a new event; if it occurred sooner, it was lumped into the preceding event. This distinction allowed us to differentiate short, momentary interactions (e.g., a few seconds of preening) from prolonged or repeated episodes (e.g., preening episodes lasting multiple minutes). Here, a “dyad” refers to an unordered pair (a,b) where interaction (a,b) and (b, a) are the same, whereas a *directional dyad* (a→b) explicitly indicates that *a* was the sender and *b* the receiver, hence (a,b) and (b, a) are different.

We scored the occurrence of behaviors (detailed in Table 1) within four primary categories—affiliative, aggressive, display, and copulatory. aggressive behaviors were further subdivided into 'courtship aggression' and 'non-courtship aggression.' Courtship aggression is defined as a female showing aggression toward a male who was displaying or courting her just before the aggressive interaction. All other aggressive behaviors fell into the non-courtship aggression category. For higher-order analyses, we combined all behaviors in each category into a single count, weighting higher-intensity behaviors as 2 instead of 1 (e.g., in the affiliative behavior category, preening has a weight of 1, and allofeeding has a weight of 2). These totals were then used to generate the interaction matrix for each behavior category.

Table 1: *Ethogram with a description of each behavioral interaction.*

Behavior Interaction Category	**Behavior**	**Description**	**Weight**
Non-Courtship Aggression	Aggression-Close	A bird showed aggression toward another bird by lunging at it with its beak, or by attacking with its beak in any other possible way.	2
	Mutual Fight	Both birds showed aggression toward each other in a quick series of attacks. The winner or "sender" in this interaction was the individual that moved away or backed out of the fight first. If this was not clear by the previous criterion, the initiator of the fight was considered the winner or sender.	2
	Chase away	A bird either threatened another bird with an open beak or moved quickly toward another bird, causing the other bird to fly away or move far along the perch.	1
Courtship Aggression	Courtship Aggression	A female showed aggression toward a male who was displaying or courting her just before the aggressive encounter.	1
Courtship	Mount	A bird mounted another bird and stayed in that position for a few seconds.	2
	Attempted Mount	A bird attempted to mount another bird, but the other bird moved away, preventing a successful mount.	1
Affiliative	Preen	A bird preened the feathers of another bird with its beak.	1
	Allofeed	A bird regurgitated food into the mouth of another bird.	2
Display	Display	An individual displayed to another individual. Here, a display is defined as a sequence of body movements detailed in Wewhare and Krishnan (2024).	1
	Mutual Display	Both individuals displayed to each other simultaneously. The bird that initiated the first display was considered the sender.	1

2.3.1.2 Creating networks

From the weekly behavioral data, we summed interactions by category to produce a weekly interaction matrix (rows = sender, columns = receiver). These matrices are

shown as heat maps in Supplementary Figure 1a–d (because these social-interaction matrices underpin the rest of our analyses, we include them in the appendix as heat maps). We also visualized these interaction matrices as directed networks (Figure 1).

2.3.1.3 General methods for null model analyses

Most of our analyses examined whether network structures or association values were larger or smaller than expected by chance. Traditional “node-level permutations” in animal social network analysis often produce artifacts, leading to possible Type I or II errors. Recent studies instead suggest a data-stream permutation approach, which preserves the underlying distributions of events (Croft et al., 2011; Farine & Carter, 2022). Therefore, for each day’s observation table, we randomly swapped the sender among events of the same behavior category and the same sex pairing (m–m, f–f, m–f, f–m), disallowing swaps that would create “self” interactions. We repeated these swaps 10 times per day of data to ensure thorough randomization. Each day’s permuted dataset was then aggregated to form a new “permuted” version of our weekly data. We repeated this entire process to generate 1,000 independent permuted datasets. Next, to infer the significance of a given metric (e.g., coefficient of variation or correlation coefficient as detailed below), we compared our observed value to the distribution of values generated by these 1,000 permutations. The p-value was the proportion of permuted values more extreme than observed, and we also reported the *z*-score of our observed value relative to the null distribution, as an effect-size measure.

This approach allowed us to assess how likely the observed patterns would be if interactions were randomly reassigned among individuals, while still matching each day’s overall activity levels, sex composition, and behavioral categories. Because the permutation test preserves most of the underlying data structure and only randomizes which individuals interact with whom, it simulates a scenario in which individuals have no specific interaction preferences. They still perform the same number of interactions for each behavioral category, but with different partners, thereby ensuring that any deviations from this null model arise from genuine

individual preferences. Consequently, this conservative approach yields robust results even with smaller sample sizes (Farine & Carter, 2022).

2.3.1.4 Non-randomness in social interactions

With our analytical framework in place, we first examined whether interactions between individuals across the network differed from chance expectations, which would indicate that certain individuals or dyads preferentially interacted with each other. This analysis was repeated each week and for each behavioral category, including subsets divided by sex (male-male, female-female, and male-female). To measure the non-randomness of interactions, we used the coefficient of variation (CV), defined as variance divided by the mean of the interaction matrix's cell values. Under purely random interactions, the CV is expected to be low because the matrix entries are relatively uniform. By contrast, strong preferences inflate the variance, producing a higher CV. We compared the observed CV against the distribution obtained from our data-stream permutation method. If the observed CV lay outside this null distribution, we concluded that the network exhibited significant non-randomness (p-value) and reported the CV's z-score as an effect-size measure.

2.3.1.5 Scale of data aggregation

Here, we evaluated whether a 1-week time frame provided a sufficiently stable representation of the social interaction network, by comparing correlation values over progressively larger time windows (1 day, 2 days, 1 week, and 2 weeks). Because our acoustic data were already collected on a weekly schedule, we aimed to check whether weekly aggregation was suitable to quantify metrics within the social behavior data. Using Spearman's rank correlation (chosen for its more flexible, non-linear assumptions), we computed correlation coefficients between adjacent time windows under each of the four aggregation scales. We anticipated that if the chosen window were too short, interactions would appear incomplete, yielding a low correlation from window to window. Correlation values would initially increase as the window widened because the chosen time window better captured each dyad's interactions. However, if the window became excessively long, temporal drift would dilute the continuity of interactions, causing correlations to decline. Observing that

correlations plateaued at 1 week indicated that weekly aggregation was both comprehensive enough to capture each dyad's interactions and not so extensive that it blurred changes over time.

2.3.1.6 Drift in social network structure

Next, we examined whether the network "drifted" over time by quantifying correlations between daily interaction matrices that were 1, 2, or 3 days apart, making sure to choose days within the same week. We then repeated this analysis across weekly intervals of 1, 2, or 3 weeks, separately for each behavior category. If drift were significant across days within a week, correlations would rapidly decrease at longer separations. If drift or changes in interactions over the scale of days were not apparent, this strengthened the rationale for aggregating data by week.

2.3.1.7 Reciprocity or dyadic correlation

Examining the stability of the social interaction network provided a launching point to examine the structure and hierarchy of the network. To do this, we first examined whether dyadic interactions exhibited directionality, namely, whether one individual's behavior in a pair influenced the other's response or whether both actions were independent. Directionality could manifest as either positive or negative reciprocity. For example, in a dominance hierarchy, frequent aggression by a higher-ranked bird toward a lower-ranked bird may be negatively correlated with aggression in the reverse direction (the lower-ranked bird rarely retaliates). Conversely, behaviors indicative of social bonding may show positive reciprocity, with both birds responding similarly to each other's actions.

Because courtship aggression (i.e., a female's immediate aggressive response to a male's attempt to display or mount) differs fundamentally from broader social or dominance aggression, we excluded it from these analyses. We thus defined our final "non-courtship aggression" category to capture dominance-related aggression. All data were aggregated at the weekly level, and we then combined the resulting weekly data from all four weeks to test for overall reciprocity effects. To ensure each pair's interactions were adequately represented, we included only dyads that had

accumulated at least five events in the relevant behavior category. Each data subset also had to contain at least five valid dyads before we performed statistical tests or plotted data. For male-female dyads, we plotted the male's values on the x-axis and the female's on the y-axis, then plotted a best-fit line to the observed data and calculated Spearman's rank correlation coefficient (rho). To determine whether the observed correlation was more or less than expected by chance, we compared it against a null distribution of correlation coefficients derived from previously generated permuted datasets. The observed p-value indicated how often the null distribution surpassed the observed rho value.

2.3.1.8 Examining the dominance hierarchy and calculating rank

Next, we used three established methods to calculate dominance relationships among all individuals, doing so separately for each week. Specifically, we calculated:

1. *David's Rank (DR)*

   David's Score incorporates both an individual bird's direct win-loss record against each opponent and the overall strength of those opponents (David, 1987). We computed a probability matrix $p_{ij} = (w_{ij})/(w_{ij} + w_{ji})$, where w_{ij} is the number of wins by bird i over bird j, and w_{ji} is j's wins over i. Each bird's "weighted wins" (w) and "weighted losses" (l) were calculated by summing p_{ij} across all opponents as rows and columns, respectively. We then included second-order terms w_2 and l_2 by multiplying the probability matrix by these w and l vectors. Finally, a bird's David's Score is given by:

   $$DS_i = (w_i + w_{2i}) - (l_i + l_{2i})$$

   Higher DS values reflected higher dominance status. We converted these scores into ordinal ranks to obtain David's Rank.

2. *Randomized Elo Rating*

This approach updated each bird's rating sequentially after every aggressive interaction, starting all birds at the same baseline (e.g., 1000). For each interaction, we calculated the winner's expected win probability E_r using the rating difference between winner and loser:

$$E_r = 1 / (1 + 10 ^{[(\text{Rating}(l) - \text{Rating}(w))/400]})$$

The winner's rating was then increased by $k \times (1 - E_r)$, while the loser's is decreased by $k \times E_r$, with k as a constant (here set to 32, standard used in most studies). To minimize bias from a single interaction sequence, we randomly shuffled the order of interactions many times (e.g., 100 repeats) and then averaged the resulting Elo scores. This yields the final rating for each bird (Sánchez-Tójar et al., 2018).

3. *Eigenvector Centrality*

We treated the win-loss matrix as a directed graph, where an edge from bird i to j had a weight representing i's wins over j. Eigenvector centrality iteratively assigns each bird a score that depends on its connections to others who themselves have high scores. Thus, each bird's centrality increased if it dominated birds that were themselves dominant. Birds with high eigenvector scores were thus embedded at the top of a strongly hierarchical network (Bonacich, 1987).

For each method, we generated weekly dominance ranks for every individual, then calculated Spearman's rank correlations between ranks derived using different methods (e.g., DS versus Elo). Provided all three approaches produced broadly consistent hierarchies, we retained David's Score for subsequent analyses, given that it is relatively straightforward yet robust (Sánchez-Tójar et al., 2018).

2.3.1.9 Sex differences in the dominance hierarchy

In many animal societies, dominance hierarchies are sex-segregated, with one sex consistently outranking the other regardless of individual variation (Ang & Manica, 2010; Paoli et al., 2006). We compared dominance ranks between males and females using David's Score to determine whether this was the case in budgerigars. A higher or more positive score indicated a higher rank. We performed a Mann–Whitney U test on the raw scores and the ranks derived from them to test for significant differences between sexes.

2.3.1.10 Temporal changes in dominance rank

In addition to tracking changes in dominance ranks over time, we quantified the stability of dominance over time by calculating the intraclass correlation coefficient ICC(1), which measures the proportion of variance in rankings of individuals across weeks explained by individual identity to the total variance in the data (Shrout & Fleiss, 1979). A high ICC indicated that dominance ranks remained stable, whereas a low ICC suggested that ranks fluctuated significantly. We used the `*intraclass_corr*` function from the *Pingouin* library to compute this metric (Revelle, 2007). ICC(1) corresponds to a one-way random-effects model, making it equivalent to the repeatability measure commonly used in the *rtpr* package in R.

2.3.1.11 Relationship between rank difference and win probability

To assess the rigidity and structure of the dominance hierarchy, we quantified three key parameters: (1) steepness (De Vries et al., 2006), (2) linearity using triangle transitivity (Ttri) (Shizuka & McDonald, 2012), and de Vries's h (h') (De Vries, 1995), and (3) the relationship between win probability and rank difference.

Traditionally, linearity has been the focus when describing dominance hierarchies, capturing whether dominance relations form a transitive chain with minimal intransitive loops. We report two measures of linearity: triangle transitivity (Ttri), which evaluates the proportion of observed triads (A, B, C) that exhibit consistent dominance (e.g., if A > B and B > C, then A > C), and de Vries's h′, which modifies

the older Landau's h to handle missing dyads by randomly assigning winners and losers in unobserved pairs. Ttri is more reliable for sparse datasets because it simply excludes triads with missing data. However, as h′ has been widely used in past literature, we included both for comparability (De Vries et al., 2006; Hobson et al., 2014; Sánchez-Tójar et al., 2018).

In contrast, steepness focuses on the magnitude of differences in success between adjacent ranks. A steep (despotic) hierarchy exhibits large gaps in dominance success from one rank to the next; a shallow (egalitarian) hierarchy shows only subtle differences, even if it is still linear (De Vries et al., 2006; Shizuka & McDonald, 2012). Here, we measured steepness by taking the slope of the regression between each individual's normalized David's Score and its David's rank: a slope near zero indicates a shallow hierarchy, whereas values closer to 1 indicate a steep hierarchy.

We assessed statistical significance for Ttri, h′, and steepness using their respective permutation or randomization tests, as outlined in each original paper (De Vries, 1995; De Vries et al., 2006; Shizuka & McDonald, 2012). However, we caution against interpreting nonsignificant results as evidence for an absence of hierarchy, because these metrics were devised primarily with steep hierarchies (e.g., in chickens *(Gallus gallus domesticus)* or chimpanzees (*Pan troglodytes*)) in mind. In shallower systems (for example, monk parakeets (*Myiopsitta monachus*)) p-values often appear nonsignificant even when clear dominance relationships exist upon closer inspection (Hobson et al., 2014; Hobson & DeDeo, 2015).

Given the limitations of steepness and linearity measures when facing shallow hierarchies, we also examined the relationship between win probability and rank difference as a more intuitive visualization of hierarchical structure. Specifically, we binned interaction outcomes based on the rank difference between the two individuals and plotted the average win probability of the higher-ranked bird in each bin. This approach, recently recommended for dominance analyses, highlights how quickly (or slowly) win probability increases with increasing rank difference (Sánchez-Tójar et al., 2018). These analyses were conducted separately for each sex for each weekly dataset, as well as on the pooled data across all weeks,

allowing us to capture both short-term fluctuations and overall trends in dominance relationships.

2.3.1.12 Correlation between different behavior categories

To infer possible functional overlaps or distinctions between behavioral categories, we investigated whether interactions in one category were correlated with or independent of those in another. Each directional dyad served as our unit of analysis, given that behaviors could differ depending on the initiating (A→B) versus the receiving individual (B→A). We aggregated data on a weekly basis, for each dyad for each category. We then combined data across all 4 weeks for the final analysis. We focused on three principal behavioral categories and compared their dyad-level frequencies using null models as discussed earlier (e.g., Category X vs. Category Y). We used Spearman's rank correlation (rho) to quantify each association and derived p-values by comparing the observed rho to that obtained from 1,000 permuted datasets. For a realistic and fair comparison, we retained only directional dyads with at least five events across both categories being compared in total, and each data subset also had to contain at least five data points overall for the final analysis to be conducted for the subset. This ensured that our correlation estimates were based on sufficiently representative observations.

2.3.1.13 Male-male displays as an appeasement gesture

To examine whether male-male displays served as appeasement gestures following an aggressive interaction, we counted display events from either individual within each male-male dyad within a fixed time window (30, 60, 90, 120, or 300 s) following an aggressive act within that dyad. To distinguish patterns, we tracked separate counts of whether the display was performed by the aggressor or by the recipient. If a second aggressive act occurred before the window ended, we halted the count at that point. We contrasted the observed display counts against an expected distribution of counts derived from 1,000 randomized datasets, where the order of events was permuted within a single day's observation. We then calculated a z-score reflecting the difference between observed and expected frequencies. Each time window yielded a corresponding p-value, and we plotted z-scores (y-axis) versus the

time window (x-axis). We marked non-significant time windows with circles and annotated the actual display counts, indicating how frequently these hypothesized appeasement gestures occurred in our dataset.

2.3.2 Syntax similarity

After examining the structure of the dominance hierarchy and correlations of behavioral events to each other, we next examined how these phenomena were reflected in vocal similarity, using both metrics of spectral similarity and syntactic structure. For each week's data for each male, we selected 10 warble sequences (hereafter "warbles"). Following previous studies, we defined a warble as a sequence of notes longer than 1 second and containing more than three distinct notes, bounded by inter-note gaps of no more than 1 second. When choosing warbles from a day's recording, we selected warbles occurring as far apart in time as possible to minimize the effects of one sequence on an adjacent one, and to capture as much variation in the repertoire as possible. Although females do not typically vocalize when a male sings, they occasionally utter a few call notes (Madabhushi, Wewhare, et al., 2023). To ensure that our warbles contained only the male's notes, we cross-checked each bout with its accompanying video and confirmed the absence of female vocalizations during the sequences we analyzed. We visualized the raw audio in Raven Pro 1.6.5 (Cornell Laboratory of Ornithology, Ithaca, NY, USA) (2024), using consistent spectrogram settings (brightness = 45, contrast = 60, window length = 512) and drew boxes around each individual note with the cursor tools, assigning a note type according to a previously established scheme (categories were: 'a', alarm call-like elements; 'b', contact call-like elements or frequency modulated sounds; 'c', long harmonic calls or harmonic sounds with a duration greater than 100 ms; 'd', short harmonic calls or harmonic sounds with duration less than 100 ms; 'e', 'noisy' calls that are broadband and non-harmonic; 'f', clicks or extremely short broadband calls, and 'm', or 'soft calls') (Madabhushi, Wewhare, et al., 2023; Wewhare & Krishnan, 2024). We also incorporated silence into the sequences, again following previously established protocols: if a gap exceeded 0.5 s seconds (half the 1-s threshold for warble termination), we inserted a "silence" element '0' (Bhat et al., 2022). To verify the objectivity of our classification,

two annotators independently labeled the same warbles, and we measured inter-annotator agreement using the Levenshtein distance on the resulting note strings. Over 10 warble sequences, average agreement was 84.5 %, which confirmed the robustness and objectivity of our classification scheme. Overall, we aimed to annotate 10 warbles per male per week. However, in the final (fourth) week, male SB did not produce any warbles (see above), so no data were available for this individual during that period. We compared every individual's warble repertoire both within and across weeks of data. Our basic comparison unit was an individual's weekly repertoire, which was then evaluated against every other individual–week repertoire combination to form a pairwise distance matrix. This matrix was computed using multiple metrics designed to capture different facets of syntactic structure.

2.3.2.1 Measures of syntactic similarity

1. *Local alignment–based distance:* Because our warble sequences varied greatly in length, aligning them in full to examine similarity could be misleading. Instead, we sought to identify shorter regions of exact similarity and emphasize those. Local alignment methods, widely used in bioinformatics to identify regions of similarity in DNA or amino-acid sequences (Lesk, 2014; Mount, 2001), are designed precisely for such scenarios, in contrast to global alignment or simpler edit-distance metrics like Levenshtein distance. Levenshtein distance (Levenshtein, 1966), for instance, measures how many single-character insertions, deletions, or substitutions are required to transform one string into another. As an example, "bein" can be converted to "pin" by substituting 'b' with 'p' and deleting 'e,' yielding an edit distance of two. Although effective for broad comparisons, Levenshtein's penalization extends across the entire length of both sequences, making it less suitable for comparing shorter regions of strong similarity embedded within longer, more variable warbles. We thus adopted the Smith-Waterman algorithm (Smith & Waterman, 1981), which uses dynamic programming to find the best local alignment rather than forcing an alignment of complete sequences. Each cell in the dynamic programming matrix encodes the optimal local alignment score for the prefixes of two sequences, penalizing gaps (insertions or deletions) and mismatches but stopping when a similarity region ends, rather than

continuing through non-similar trailing segments. We assigned +1 for matching note types and −2 for both mismatches and gaps. This relatively high penalty discourages partial or approximate alignments, focusing attention on strictly matching segments. Because spurious matches can arise when very short warbles are aligned to much longer ones, we included a length penalty by dividing the final alignment score by the natural logarithm of the sum of the combined sequence length for the two sequences. A high alignment score thus reflects a genuinely similar (and appropriately scaled) subsequence. To obtain a repertoire-level comparison, we designated one individual's weekly warble repertoire as the "focal list" and another as the "matched list," calculating alignment scores for every pair of warbles between them. The average of these scores from all pair-wise comparisons formed the similarity measure for the focal repertoire against the matched repertoire. However, this similarity might differ if the roles of "focal" and "matched" were reversed, so we computed the alignment in both directions and then averaged those two scores for a final, symmetrical similarity metric. Because our other syntax metrics yielded distances, we inverted the similarity measure by taking its negative value, ensuring that higher similarity corresponds to lower distance in all metrics.

2. *Tri-gram distance:* We applied an *n*-grams approach—specifically, tri-grams (n = 3) to further examine syntactic similarity. Many animal vocal sequences display significant repetition, and by capturing sub-sequences of length 3, we captured local structural patterns without overly fragmenting the data (Kershenbaum et al., 2016; Kershenbaum & Garland, 2015). There are (L−n+1) sub-sequences of length n for a sequence of length L. For instance, in the warble segment "ABBAC" (length 5), there are (5−3+1)=3 distinct tri-grams: "ABB," "BBA," and "BAC." If the warble repertoire has C distinct note types, then there are C^n theoretically possible n-grams. We represented each sequence as a *feature vector* of tri-gram frequencies, normalized by the total tri-gram count to yield relative proportions. We computed the *Euclidean distance* between tri-gram proportion vectors to compare two sequences (or aggregated repertoires for individuals). A smaller distance implied more similar tri-gram distributions, indicating that both sequences contained similar

repeated triplets of notes (Kershenbaum & Garland, 2015). Although the tri-grams possess underlying Markov assumptions about the underlying process of sequence generation, we used them here simply to measure similarities and not to assess the underlying process (Kershenbaum et al., 2014).

3. *Co-occurrence counts:* We constructed a count matrix counting the number of times note B occurred within d steps of note A, to examine how often particular note types appeared near one another. For each occurrence of A in a sequence, we counted the number of times each note B appeared within a distance of d notes, summing this value across all occurrences of A (Bhat et al., 2022; Madabhushi, Bhat, et al., 2023). This process produced a $|\text{note-types}| \times |\text{note-types}|$ matrix for each sequence. When generating a repertoire-level matrix for an individual, we simply added the corresponding matrices from all warbles in that repertoire. We then normalized each matrix to its overall sum, making them comparable across different repertoires. Finally, we measured Euclidean distances between these normalized matrices to obtain a co-occurrence count-based syntax distance (Bhat et al., 2022).

4. *Co-occurrence ratio:* We also computed a co-occurrence *ratio* by dividing each observed count by its expected value under a randomized arrangement of notes. We derived this expected count by taking the average of 100 randomized count matrices. Each of these count matrices was derived from sequences where the order of notes within the sequence was randomly permuted. Dividing the observed by the expected gave an odds ratio. We then took the natural logarithm of each odds ratio, ensuring that under- and over-representation of pairs relative to chance were treated symmetrically (Wewhare & Krishnan, 2024). Finally, we obtained distances by computing the Euclidean distance between the log-odds ratio matrices of two repertoires.

5. *Combined syntax metric:* Because we found the tri-gram and co-occurrence counts to be highly correlated, our final comprehensive syntax distance metric was the average of three measures. Before averaging, we normalized the

values within each distance matrix between 0 and 1 to ensure all metrics were given equal weight:

1. Local alignment
2. Tri-gram distance
3. Co-occurrence ratio

These resulting repertoire-level distance matrices were employed in subsequent analyses linking warble syntax similarity to social structure. Because these metrics perform best with larger datasets, we prioritized repertoire-level comparisons (i.e., aggregating multiple warbles per individual) rather than single-sequence comparisons, when linking the acoustic network to the social network.

2.3.2.2 Individual sequence analyses

To examine finer nuances of each bird's vocal repertoire, we shifted our analysis to the single-sequence level, creating a new set of distance matrices in which every warble was compared pairwise with all other warbles (both from the same individual and from others). Because certain syntax metrics are most effective for repertoire-level comparisons, we restricted ourselves to two measures: local alignment (Smith-Waterman) and tri-gram distance, as both adapt well to single-sequence assessments. Local alignment is designed to find the longest region of similarity between two sequences, whereas tri-gram distance, as a summary metric of the sequence, is the least computationally complex and the simplest among all other metrics. Thus, it should provide comparably more robust results even with less data. We then calculated a combined distance metric by averaging these two measures. Before averaging, we normalized the values within each distance matrix between 0 and 1. Once this pairwise distance matrix was computed, we visualized it using a two-dimensional UMAP projection. UMAP attempts to preserve the relative distance relationships among points when reducing dimensionality, making it effective for capturing both local and global structures. This approach has been shown to outperform alternatives such as NMDS in maintaining these relationships during the projection (Armstrong et al., 2022). This approach enabled us to examine

individual and temporal signatures in warble songs, by ordinating them in syntactic space.

2.3.2.3 Individual signatures

If warbles contained individual-specific signatures, we would expect within-individual (intra-individual) distances to be smaller than between-individual (inter-individual) distances. We tested this using a permutation approach. First, we computed the mean intra-individual distance and the mean inter-individual distance across all warbles and calculated the difference between them. We then permuted the labels, assigning each warble to individuals (i.e., randomizing bird identity), recalculated the mean intra- and inter-individual distances, and computed the difference. Repeating this process 1,000 times produced a null distribution of distance differences. Comparing our observed difference to this null distribution yielded a p-value and z-score, indicating whether the observed pattern (lower intra- than inter-individual distance) was significantly stronger than expected by chance.

2.3.2.4 Week-level signatures

Next, we examined whether warbles recorded in the same week (regardless of which individual produced them) were more similar than to warbles from different weeks, i.e. whether there was temporal drift in syntax during the period of data collection. We repeated the above permutation tests but randomized week labels instead of individual labels, computing the difference between intra-week and inter-week distances.

2.3.2.5 Signature of drift within individual repertoires

Finally, to investigate how each individual's warbles changed over the course of the study, we focused on a single bird's repertoire across all weeks. We again performed a UMAP projection for that individual alone, and tested for a week-level signature in the same manner as above (permuting only the week labels within that individual's data). A significant effect would suggest that a single bird's vocal sequences drifted

or shifted syntactically over time, indicated by smaller intra-week distances than inter-week distances.

2.3.3 Spectral analysis of contact call note structures

Our last major set of analyses focused on the time-frequency structure of b notes, which are contact call-like elements embedded within warbles (Farabaugh et al., 1992). These frequency-modulated calls are known to encode both individual and group identity in budgerigars and other psittacine species (Manabe et al., 2008; Tu et al., 2011). Given their potential role in social communication, we specifically examined variation in call structure across individuals. Each bird produces a small repertoire of contact call subtypes, some of which may be unique to that individual, whereas others may be shared with close social associates (Tu et al., 2011; Zhao & Goldberg, 2023). Previous studies have demonstrated that budgerigars learn the contact calls of other individuals in their flock, and reproduce them to either attract attention or signal affiliation (Dahlin et al., 2014). We analyzed contact call-like elements—notes within warbles that spectrally resemble contact calls given in isolation. First, we extracted b notes from the same warbles previously used for syntax analysis to investigate this variation. From these warbles, we sampled up to 100 b notes per individual per week, although in some cases, fewer calls were available: FL produced only 71 calls in Week 1, SR had 91 calls in Week 3, and SB did not vocalize in Week 4, resulting in no data for that period. For spectrogram visualization, we used the Python audio processing library *Librosa* (McFee et al., 2015) with a Hanning window, setting the window length to 256 and the hop length to 64. These parameters ensured a balance between spectral and temporal resolution, enabling clear identification and examination of call structure.

2.3.3.1 Quantifying call similarity using Dynamic Time Warping

Following the approach used in the previous syntax analysis, we constructed a pairwise distance matrix to quantify measures of similarity between all extracted b notes. We applied a modified version of Dynamic Time Warping (DTW) to measure spectral or acoustic distances between calls. DTW improves upon Spectrogram

Cross-Correlation (SPCC) by addressing a key limitation of SPCC: calls with highly similar frequency modulation patterns may yield low correlation values if one call is slightly stretched or compressed in time. DTW overcomes this by using dynamic programming to assess whether similarity increases when calls are temporally adjusted (Kaprykowsky & Rodet, 2006). In our implementation, we used cosine distance to compare spectrogram columns, enabling DTW to maximize the similarity between corresponding time points when aligning two calls. The final DTW distance was computed as the sum of cosine distances across the most optimal alignment path. Because DTW sums over matrix columns, longer calls naturally yield higher raw distance values. To correct this, we normalized the final DTW score by dividing it by the sum of the lengths of the two spectrogram matrices being compared (Kaprykowsky & Rodet, 2006). Although this type of correction is uncommon in bioacoustics, it is widely applied in human speech processing, where the same word may be spoken at different tempos while retaining its characteristic structure. Our motivation for applying this correction was similar: individual identity is thought to be encoded primarily in the frequency modulation pattern rather than the absolute call duration. By normalizing for call length, we ensured that DTW distance reflected spectral similarity rather than being biased by call duration.

2.3.3.2 Clustering into call subtypes

Using the normalized DTW method, we generated a pairwise distance matrix capturing the similarity between all b note calls. To visualize the structure of these calls, we applied UMAP dimensionality reduction, projecting the high-dimensional distance matrix into two dimensions. The UMAP projection revealed distinct clusters, which was expected given that budgerigars are known to produce subtypes of b notes. To formally delineate these clusters, we applied DBSCAN clustering, a density-based algorithm particularly well-suited for identifying clusters of varying shapes. Unlike k-means, DBSCAN does not require the number of clusters to be specified a priori, making it more appropriate for this type of data (Ester et al., 1996). DBSCAN relies on two key parameters: (1) the minimum number of points required for a core cluster point, which we set to 10, and (2) the maximum distance between points for them to be considered part of the same cluster, which we set to 0.2. These

values were determined by manually adjusting the parameters to achieve the most coherent clustering of call subtypes.

2.3.3.3 Individual and week-level signature

Again, following the methodology used in the syntax analysis, we tested for individual-level signatures in the time-frequency structure of b notes. This analysis was conducted at two levels: first, by comparing distances across the entire dataset, and second, by restricting comparisons to within-cluster distances. The procedure was identical to the approach used previously, except that the DTW-based metric was used for distance calculations instead of syntax-related measures. To determine whether b notes exhibited temporal signatures of changes across weeks, we applied the same methodology used in the syntax analysis, testing whether calls within the same week were more similar than those across weeks. As with individual signatures, we conducted this analysis both across the full dataset and within clusters to ensure that any week-level effects were not obscured by broader structural differences between clusters. We examined drift within individual repertoires over time, using the same approach applied to syntax data. This analysis tracked whether an individual's calls shifted in structure across weeks of data and was assessed using DTW-based distances.

2.3.3.4 Metrics for comparison between acoustic data and the social network

To assess the sharing of call subtypes across individuals, we quantified the number of clusters shared between each pair of individuals on a weekly basis. An individual was considered a member of a given cluster if at least 20% of its calls in that week belonged to the cluster, a threshold exceeding the random expectation of 16.67% (assuming equal participation across six birds). To further examine the vocal overlap of call subtype use between individuals, we computed the total number of calls an individual produced in clusters that were also occupied by another specific individual. This count was summed across all shared clusters for each dyad using the same cluster membership criteria as before. Finally, we compared the average repertoire-level distance between individuals across weeks. This was done at two

levels: first, by averaging distances across the full dataset, and second, by considering only distances within clusters and averaging across all clusters.

2.3.4 Linking acoustic distances and social structure

Having examined both the structure of the social network and the vocal similarity network of males, we next aimed to investigate whether social interactions influence vocal similarity. Within these networks, individuals were represented as nodes, and the relationships between them (dyads) were represented as edges that encoded either behavioral interactions or acoustic similarity. We conducted our analysis at two levels: the dyadic level, where we examined the relationships between edge properties across the two networks, and the node level, where we assessed whether individual properties of one network were correlated with those of the other.

2.3.4.1 Vocal similarity between interacting dyads

At the dyadic level, we tested whether a correlation existed between a behavioral metric derived from the social network and an acoustic metric derived from the vocal similarity network. The acoustic similarity metrics fell into three broad categories: syntax-based, spectral-based, and call subtype use. All acoustic comparisons were made at the repertoire level within a given week, enabling us to compare the vocal similarity between two individuals in relation to their social interactions.

Syntax-based metrics

- Local alignment distance
- Tri-gram distance
- Co-occurrence counts
- Co-occurrence ratio
- Combined syntax metric

Spectral-based metrics of b Notes

- Mean call distance across the entire repertoire
- Mean call distance within the same clusters

Call subtype use

- Number of common clusters (clusters shared between two individuals)
- Number of calls in common clusters (the sum of calls within shared clusters)

We ensured symmetry where necessary for all metrics that were extracted from matrices. If a metric was not inherently symmetric, it was made symmetric by averaging the two directional values for a dyad, thereby ensuring that the unit of analysis was the dyad rather than a directional interaction. This transformation ensured that all dyadic comparisons reflected mutual relationships rather than directional influence.

2.3.4.1 Behavioral metrics and hypothesis testing

To investigate the relationship between social interactions and vocal similarity, we selected behavioral metrics that represented different aspects of male-male interactions. Because acoustic data (i.e. warble song) were available only for males, our analysis was restricted to male-male dyads. To ensure consistency with the acoustic similarity metrics, we symmetrized all behavioral measures so that they reflected mutual rather than directional relationships as discussed above.

The behavioral metrics derived from the social network included:

1. *Number of display interactions* : quantified by how often one male displayed to another within the given data-clubbing window (i.e. one week).
2. *Number of displays to the same female(s)*: for a given male dyad (A, B), we identified all common females (C) whom both males had interactions with and summed the number of displays each male directed toward those shared females.
3. *Number of aggressive interactions:* measuring how frequently one male engaged in aggressive interactions with another.
4. *Dominance rank difference*: which was calculated based on David's Score dominance hierarchy.

Our basic unit of analysis was one week, across which we calculated dyadic behavioral and acoustic similarity values. We pooled data across all weeks in our final hypothesis testing to assess broader patterns. We systematically examined all possible pairwise combinations of behavioral and acoustic metrics to test for social and acoustic similarity correlations. We employed two complementary statistical approaches:

1. Spearman's correlation coefficient (rho), quantified the monotonic association between behavioral and acoustic similarity.
2. Linear regression analysis focused on the slope of the relationship between the two variables.

We constructed a null model using our data-stream permutation approach to determine whether the observed correlations and slopes deviated from expectations under chance. The entire analysis pipeline was repeated on these permuted behavioral datasets, generating a null distribution of correlation coefficients and regression slopes for each pair of behavioral and acoustic metrics. We reported the p-value and z-score for each observed value by comparing it to the null distribution, as with all other analyses. This approach followed best-practice guidelines for hypothesis testing in animal social networks, ensuring that our conclusions accounted for the structured nature of social interaction data (Croft et al., 2011; Farine & Carter, 2022).

We measured acoustic similarity across multiple metrics for each acoustic dimension (syntax, contact call structure, and call subtype usage). To ensure robustness, we considered an effect significant only if all metrics within the same acoustic dimension indicated significance in the same direction. Next, we assessed each effect using two metrics: the slope of the regression and Spearman's rank correlation. We required that both metrics showed, at minimum, the same direction of the effect with at least one of them being significant. However, because the regression slope is more sensitive to spurious effects from outliers, we placed greater emphasis on Spearman's correlation. Therefore, in cases where both metrics indicated the same

direction but only the correlation coefficient was statistically significant, we interpreted this as evidence of a biologically relevant effect.

2.3.4.2 Node-level analyses: linking dominance rank and acoustic similarity

At the node level, we examined whether an individual's position in the dominance hierarchy influenced its position in the acoustic similarity network. Specifically, we tested whether individuals were more acoustically similar to the most dominant male than expected by chance, which would suggest that males attempt to copy the vocalizations of high-ranking individuals. To quantify an individual's position in the acoustic similarity network, we first transformed the acoustic distance matrix into a similarity matrix by taking the negative of the distance values and adding the highest absolute distance value to ensure all similarity scores were positive. We then constructed an acoustic similarity network, where edges reflected similarity values rather than distances. To measure an individual’s connectedness in this network, we calculated its weighted degree, summing the similarity values of all its connections. To enable comparisons across different networks and weeks, we normalized this value by dividing it by the highest weighted degree observed in that network. This ensured that values were comparable across weekly networks, enabling us to pool data across weeks for the final analysis.

We visualized the acoustic distance matrix, the corresponding similarity network, and the relationship between an individual's normalized weighted degree and its dominance rank. After aggregating rank and normalized weighted degree values across weeks, we computed Spearman’s correlation coefficient to quantify the relationship between dominance rank and acoustic network centrality. We used a node-level permutation approach instead of pre-node (data-stream) permutations to test whether the observed correlation deviated from chance expectations. This choice was based on two considerations. First, in a pre-node permutation of the data stream, only the identity of aggression initiators and receivers would change, whereas the total number of aggression events per individual would remain largely the same. However, because the overall frequency of aggression given and received is integral to rank calculation, such permutations would not sufficiently disrupt the dominance hierarchy, leading to a null model that was too similar to the observed

data. Our simulations confirmed this, showing that pre-node permutations preserved much of the original hierarchy structure. Secondly, a node-level shuffle was the appropriate null model because our hypothesis focused on individual node positions rather than dyadic interactions.

Therefore, we permuted individual dominance ranks, shuffling them randomly while keeping the structure of the acoustic similarity network unchanged. After each permutation, we recalculated the correlation between rank and normalized weighted degree. Repeating this process 1,000 times generated a null distribution of correlation coefficients, against which we compared our observed correlation. We reported the p-value and z-score of our observed correlation relative to the null distribution. This analysis was repeated separately for each acoustic similarity network derived from the different acoustic metrics, ensuring that our conclusions were robust across different measures of vocal similarity.

Chapter 3: Results

3.1 Social structure

In total, we recorded 16 days of behavioral observations, each day contributing one hour of focal video data (for a total of 16 analyzed hours of video). Across these sessions, we documented 3356 dyadic interactions, which we classified into four behavioral categories: affiliative (343 events), aggression (1185 events), copulatory (50 events), and display (1778 events). To visualize these behaviors in a social network context, we constructed separate weekly interaction matrices for each behavioral category (heat maps are in Supplementary Figure 1a–d). When combined across the entire study period, these interaction networks (Figure 1) illustrate how frequently each individual interacted with others.

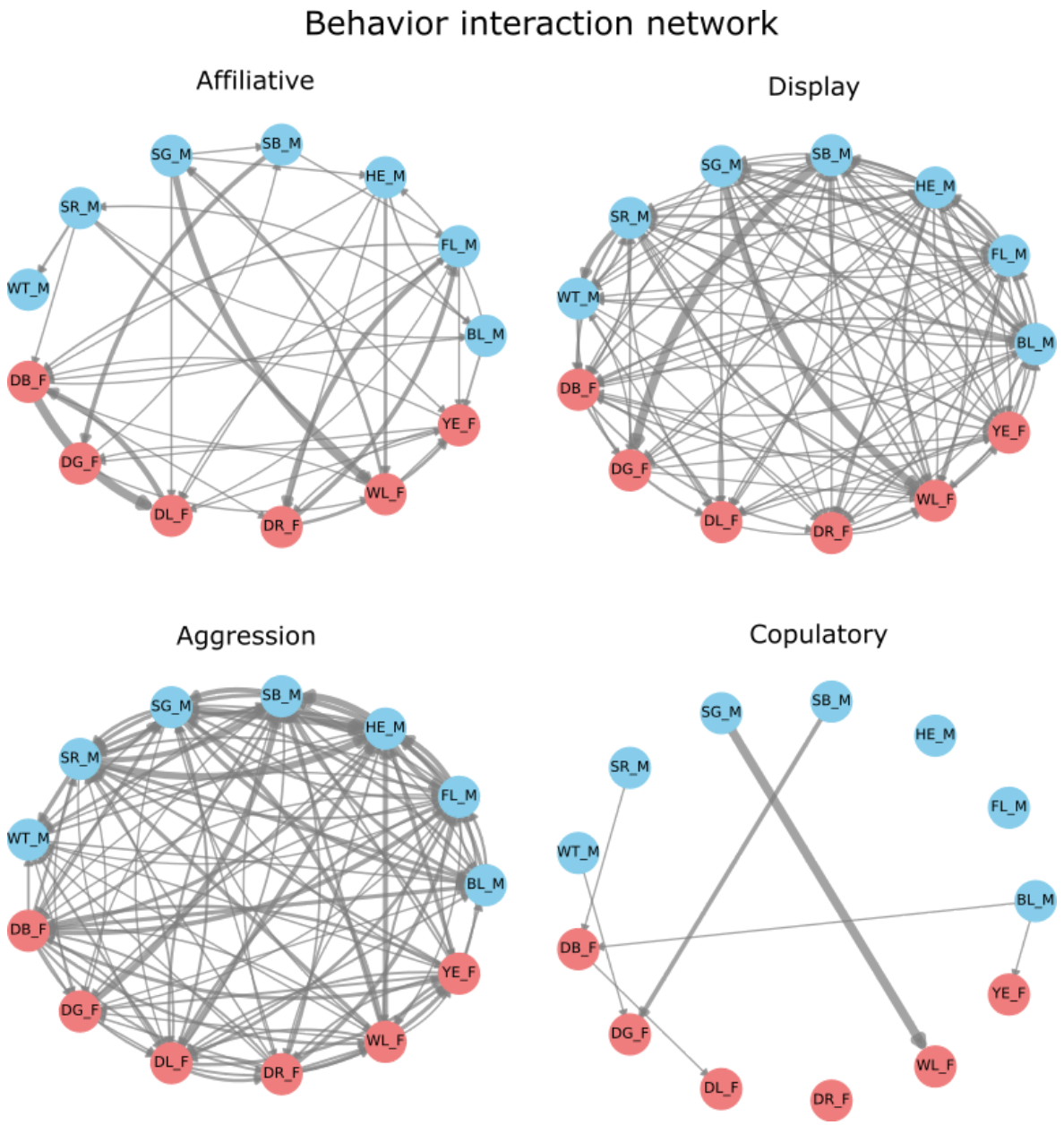


***Figure 1**: Social networks combining interaction data from all recorded weeks. Each node represents an individual bird, with females represented by pink nodes and males by blue nodes. Edge thickness is proportional to the total interaction frequency of the directional dyad, and each network's edges are normalized to its maximum value so that different categories remain visually comparable. Thus, thicker edges indicate more frequent interactions relative to other dyads in the network.*

3.1.1 Non-randomness in social interactions

First, we examined whether dyadic interactions occurred randomly, or whether certain dyads exhibited more frequent interactions than expected by chance. For each behavioral category, as well as for sex-specific subsets (male-male, female-female, and male-female), we computed the coefficient of variation (CV) of the interaction matrix (Table 2). Overall, most interaction matrices deviated significantly from randomness (high CV values, $p<0.05$).

- **Affiliative behavior**: In the combined dataset, affiliative interactions were significantly non-random for each week of data. Female-female and male-female dyads consistently showed strong preferences ($p<0.05$) in most weeks, whereas male-male interactions were less frequent overall and reached significance only in Week 2.

- **Display behavior**: All weekly interaction matrices were significantly non-random ($p<0.05$). Male-female dyads in particular exhibited higher specificity, suggesting that males selectively displayed more to certain females.

- **Aggressive behavior**: Aggression was also significantly non-random, both when considering the entire dataset and male-female dyads alone. Within male–male and female–female subsets, evidence for non-randomness was mixed (two weeks significant, two weeks non-significant). Nevertheless, CV values were higher for male-male aggression than for female-female aggression, suggesting that males directed aggression more selectively.

- **Copulatory behavior**: We did not detect significant non-randomness (possibly due to the lower sample size, n=50). However, visual inspection of the matrix (Supplementary Figure 1d) and the combined network (Figure 1) suggested some preferential behavior during courtship.

	Affiliative - Week 1	**Affiliative - Week 2**	**Affiliative - Week 3**	**Affiliative - Week 4**
All	3.36 (p<0.001)	2.98 (p<0.001)	2.75 (p<0.001)	3.72 (p<0.001)
Male-Male	1.64 (p=0.49)	1.35 (p<0.001)	nan	1.63 (p=0.35)
Female-Female	3.15 (p<0.001)	2.46 (p<0.001)	1.62 (p<0.001)	2.14 (p=0.21)
Male	2.92 (p<0.001)	3.02 (p<0.001)	2.66 (p<0.001)	2.92 (p<0.001)
	Display - Week 1	**Display - Week 2**	**Display - Week 3**	**Display - Week 4**
All	2.61 (p<0.001)	3.18 (p<0.001)	2.44 (p<0.001)	2.34 (p<0.001)
Male-Male	1.53 (p<0.001)	2.28 (p<0.001)	1.78 (p<0.001)	1.94 (p<0.001)
Female-Female	2.25 (p<0.001)	1.64 (p<0.001)	1.83 (p=0.01)	1.73 (p<0.001)
Male-Female	3.7 (p<0.001)	4.77 (p<0.001)	3.68 (p<0.001)	3.48 (p<0.001)
	Aggression - Week 1	**Aggression - Week 2**	**Aggression - Week 3**	**Aggression - Week 4**
All	1.71 (p<0.001)	2.01 (p<0.001)	1.77 (p<0.001)	1.7 (p<0.001)
Male-Male	1.42 (p=0.28)	1.64 (p=0.19)	1.47 (p=0.03)	1.46 (p<0.001)
Female-Female	1.37 (p=0.06)	1.41 (p=0.18)	1.32 (p=0.41)	1.02 (p=0.10)
Male-Female	2.11 (p<0.001)	2.67 (p<0.001)	1.96 (p<0.001)	2.61 (p<0.001)
	Courtship - Week 1	**Courtship - Week 2**	**Courtship - Week 3**	**Courtship - Week 4**
All	1.04 (p<0.001)	2.06 (p=0.10)	1.92 (p=0.10)	2.08 (p=0.49)
Male-Male				
Female-Female				
Male-Female	1.04 (p<0.001)	2.06 (p=0.09)	1.92 (p=0.09)	1.42 (p=0.51)

Table 2: *Value of CV and p-value, across interaction matrices for each behavioral category.*

3.1.2 One week is an appropriate time scale for data aggregation

To verify that combining data over a weekly time scale adequately captured the structure of the social network adequately, we examined how correlations in interaction patterns changed when we gradually increased our data-clubbing window from 1 day to 2 weeks (Figure 2a). For display and aggression, the two most frequent behaviors, the correlation coefficient between successive windows rose from approximately 0.4-0.5 at the 1-day scale to around 0.6 at the 1-week scale, indicating that a week's data is sufficient to yield a stable estimate of dyadic interactions. By contrast, affiliative and courtship behaviors are relatively rare but highly specific, and so the correlation coefficients did not show a pronounced increase with broader aggregation. Nonetheless, heat maps (Supplementary Figure 1a,d) confirmed that these infrequent behaviors still exhibited strong dyadic patterns at the weekly scale. Based on these findings, we adopted one week as the core temporal unit of analysis for all subsequent tests. This means that across the 4-week window, data for each dyad was considered for each of the 4 weeks.

3.1.3 There is no drift in the social network within a week

Next, we analyzed whether the social network was stable across weeks of data collection, or showed "drift" (i.e., a systematic change in interaction structure) over time. When comparing correlation coefficients between daily matrices with varying gaps (1–3 days apart), we did not observe a strong decline (Figure 2b), suggesting minimal day-level drift. This finding further supported our choice of aggregating data by week.

Over longer intervals, we observed a slight decrease in correlation between weekly matrices as the interval between weeks increased, suggesting a modest drift in the network structure over time. However, the number of data points available to compare matrices across larger weekly intervals was very low (only two points for a 2-week gap and one point for a 3-week gap). Due to this limited data, we did not perform a formal statistical test to confirm the significance of this observed trend. In summary, our results indicate stable social interactions at the weekly scale, with only tentative evidence for gradual structural changes across longer intervals.

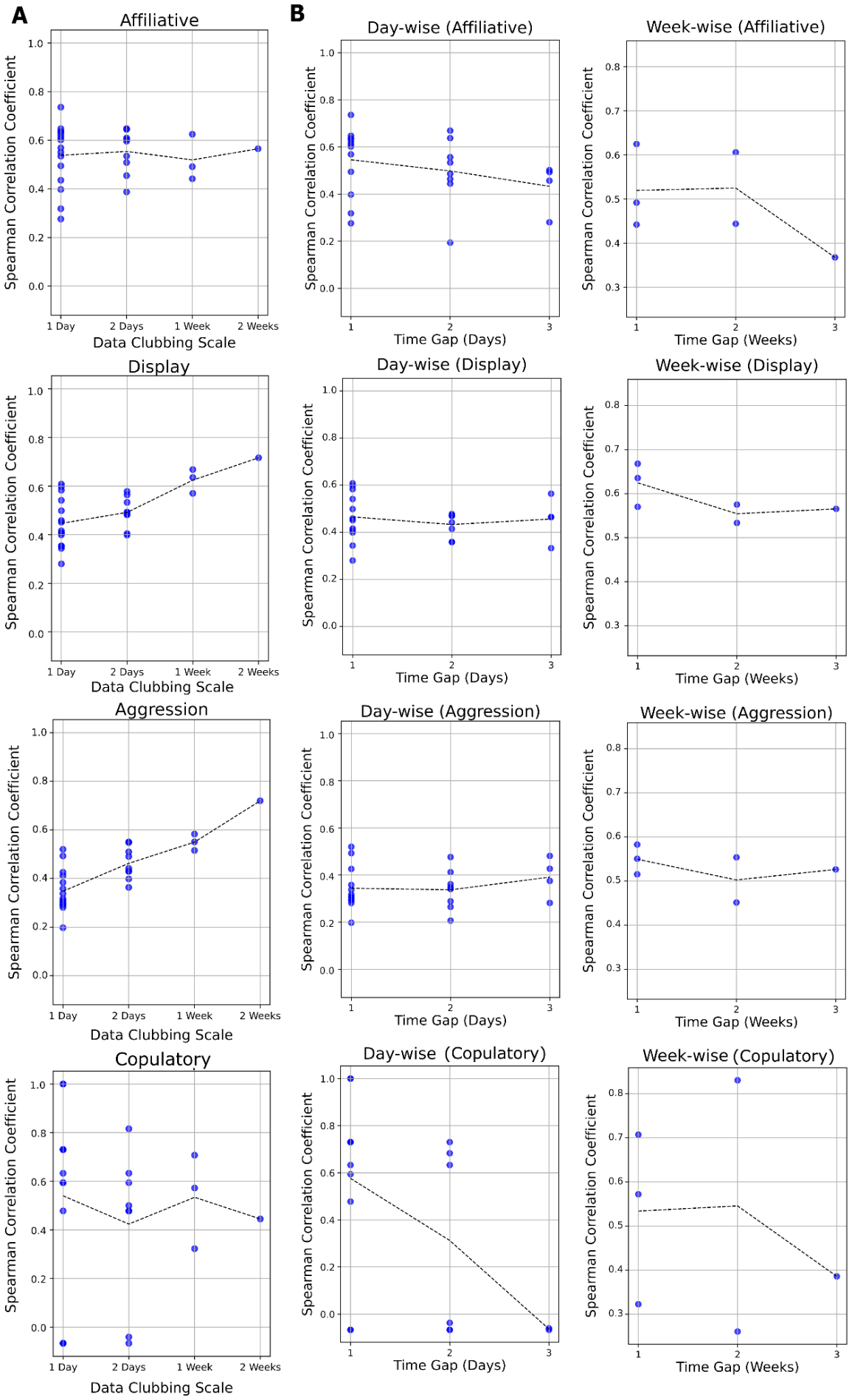

A
Affiliative
Spearman Correlation Coefficient
1 Day
2 Days
1 Week
2 Weeks
Data Clubbing Scale
B
Day-wise (Affiliative)
Time Gap (Days)
Week-wise (Affiliative)
Time Gap (Weeks)
Display
Day-wise (Display)
Week-wise (Display)
Aggression
Day-wise (Aggression)
Week-wise (Aggression)
Copulatory
Day-wise (Copulatory)
Week-wise (Copulatory)

Figure 2: (A) *This scatter plot depicts the Spearman correlation coefficient of behavioral interaction matrices in adjacent time periods, aggregated (or 'clubbed') over progressively larger windows—from 1 day up to 2 weeks. Each point represents a correlation value for a given time-window size, and the mean correlation coefficient for each window category is connected by lines. All data are shown for a single behavioral category per panel. (**B)** This scatter plot shows how the Spearman correlation coefficient of behavioral interaction matrices changes as the time gap between samples increases. Points represent correlation values at different time gaps, and the lines join mean correlation coefficients across time-lag categories. The analysis is shown on both daily and weekly scales, with each panel focusing on a different behavioral category.*

3.1.4 Negative reciprocity in aggressive interactions

We next quantified whether behaviors were reciprocated within dyads by examining correlations in the frequencies of A→B and B→A interactions (Table 3). using Spearman's rank correlation coefficient (rho) to measure reciprocity. We define positive reciprocity as a positive correlation (rho>0), i.e. higher interaction frequency in one direction predicts higher frequency in the opposite direction, and negative reciprocity as a negative correlation (rho<0). Scatterplots of observed data with best-fit lines, along with lines from randomized permutations, are shown in Figure 3.

- **Display**: Among male-male dyads, displays were positively correlated (rho=0.44, p=0.027). In contrast, female-female dyads exhibited a weak negative correlation between displays (rho=−0.06, p<0.001), suggesting that when one female displayed more frequently to another, she tended to receive fewer displays in return than expected by chance, although this effect was weak.

- **Affiliative**: Negative reciprocity (rho=−0.28, p<0.001) was observed in the overall dataset, driven primarily by male-female dyads (rho=−0.61, p=0.01).

- **Aggression**: Negative reciprocity characterized all data subsets (rho=−0.45, p<0.001 for the full dataset), including male-male (rho=−0.31, p<0.001) and female-female (rho=−0.37, p=0.023) dyads. This pattern implies that aggression from one individual in a dyad was not reciprocated by the other, pointing to an asymmetric relationship often associated with dominance hierarchies. Male-female pairs also showed a strong negative correlation (rho=−0.55) but this was not statistically significant (p=0.907).

Behavior Category	Interaction Type	Observed Correlation	P-Value
Display	all	0.008	1
Display	male-male	**0.44**	**0.027**
Display	female-female	**-0.062**	**< 0.001**
Display	male-female	-0.20	0.80
Affiliative	all	**-0.28**	**< 0.001**
Affiliative	female-female	-0.02	0.64
Affiliative	male-female	**-0.61**	**0.01**
Non-Courtship Aggression	all	**-0.45**	**< 0.001**
Non-Courtship Aggression	male-male	**-0.30**	**< 0.001**
Non-Courtship Aggression	female-female	**-0.37**	**0.02**
Non-Courtship Aggression	male-female	-0.54	0.90

Table 3: *Correlation coefficients indicating reciprocity of behaviors within dyads.*

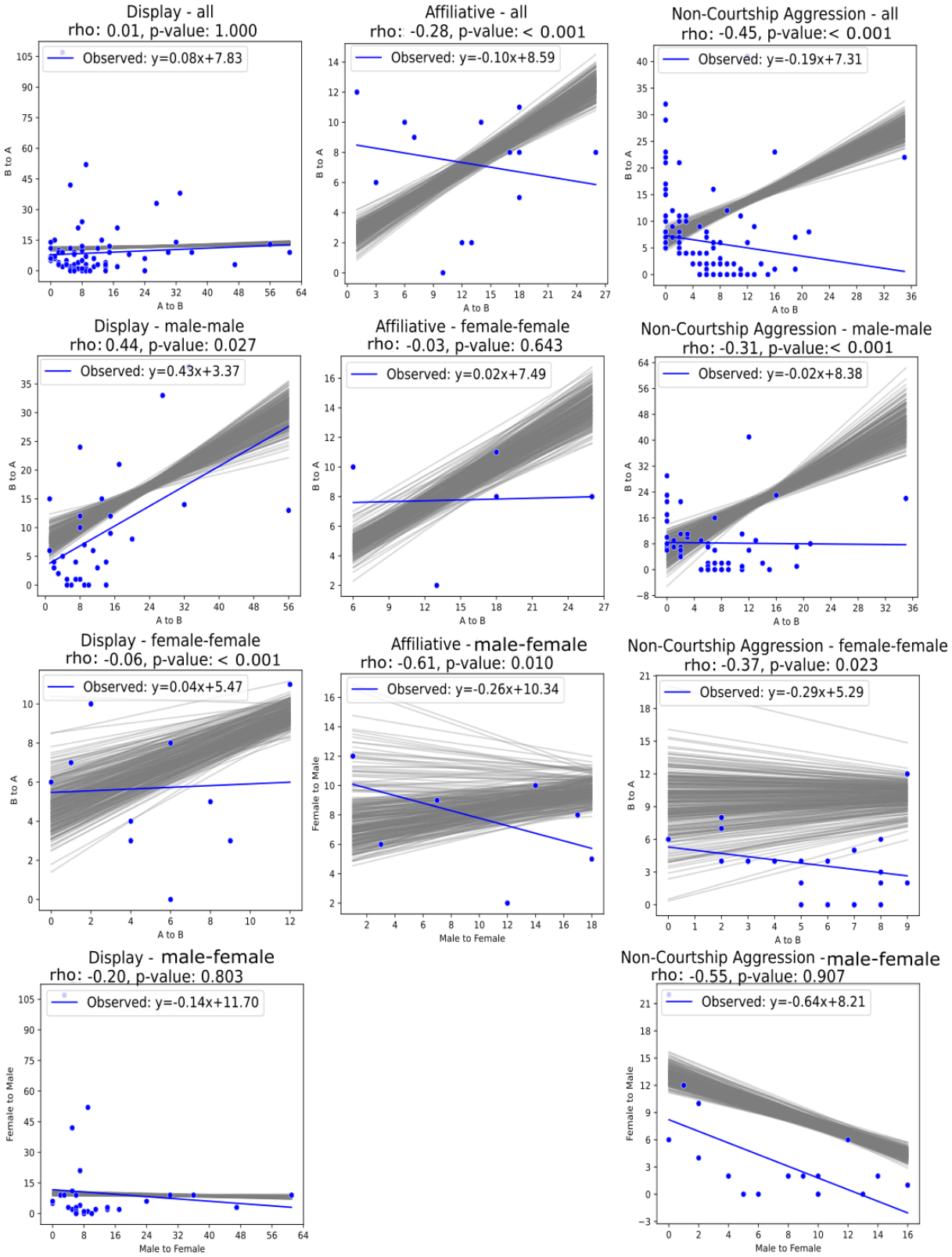


Figure 3: *Each scatter plot shows the number of interactions for both birds in a dyad (x- and y-axes) in a given week. The blue regression line is the best fit for the observed data; the gray lines are the best fit for permuted (randomized) datasets. This analysis is repeated for each behavioral category (and separately for each sex, where indicated). The plot title notes which subset of data is being shown (e.g. a specific behavior, or a particular sex*

combination), the observed Spearman correlation coefficient (rho), and the p-value from the permutation test.

3.1.5 Budgerigar dominance hierarchies are stable but not steep

Our findings of negative reciprocity in aggressive behavior led us to investigate whether a dominance hierarchy exists in budgerigars. We estimated individual dominance ranks using three standard methods: Eigenvector centrality, randomized Elo rating, and David's Score (DS). Correlations among rank estimates were high (Figure 4), with pairwise average correlation values of 0.80 (Eigenvector vs. David's), 0.76 (Eigenvector vs. Elo), and 0.90 (David's vs. Elo). Because all three methods consistently generated similar hierarchies, we selected David's score for subsequent analyses, given its simplicity and established reliability (Sánchez-Tójar et al., 2018).

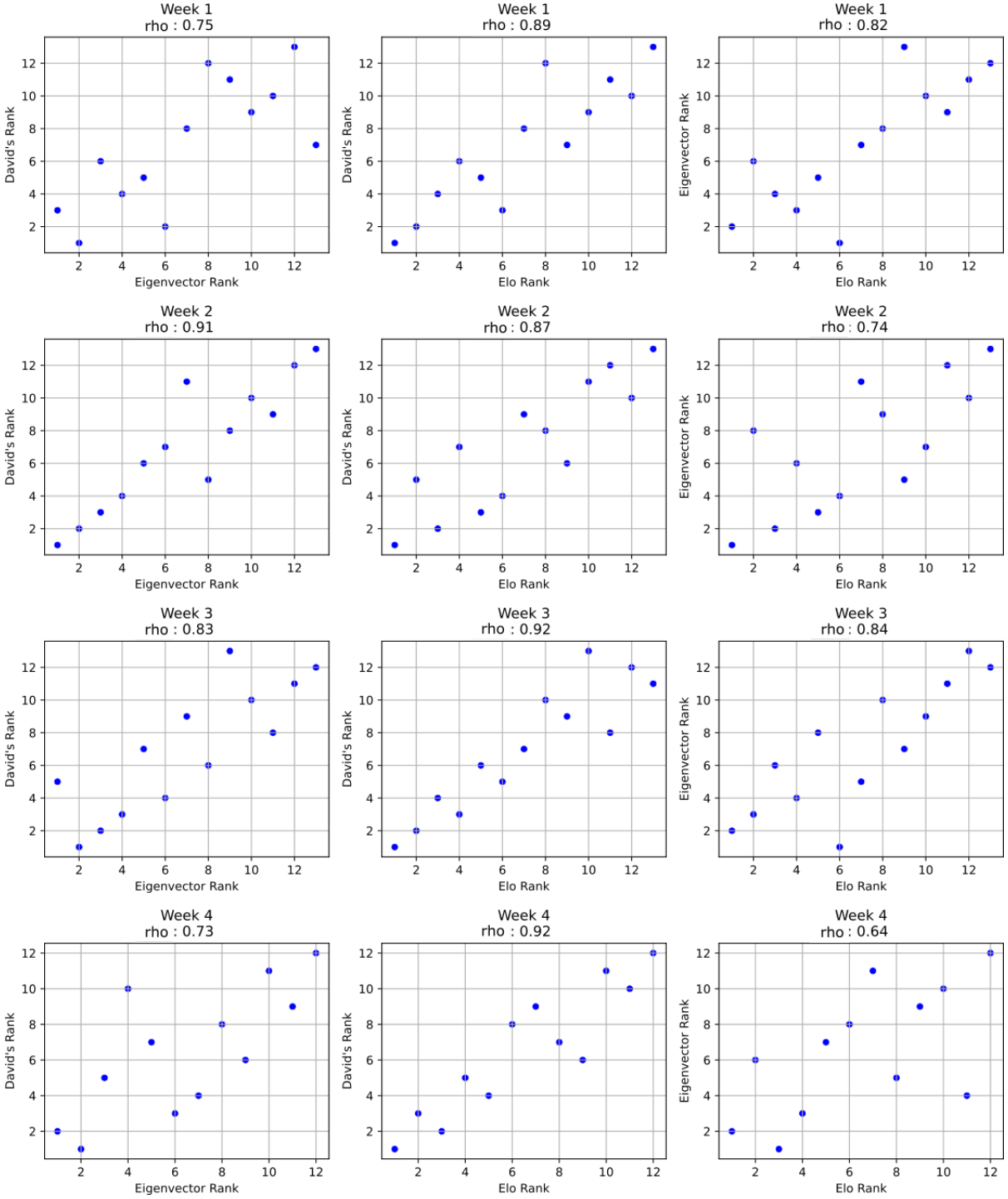


***Figure 4**: These scatter plots compare individual dominance ranks obtained from two different ranking methods. Each plot represents a pair of methods (e.g. Method A versus Method B) across all weeks. The plot titles include the Spearman correlation coefficient for the two sets of ranks.*

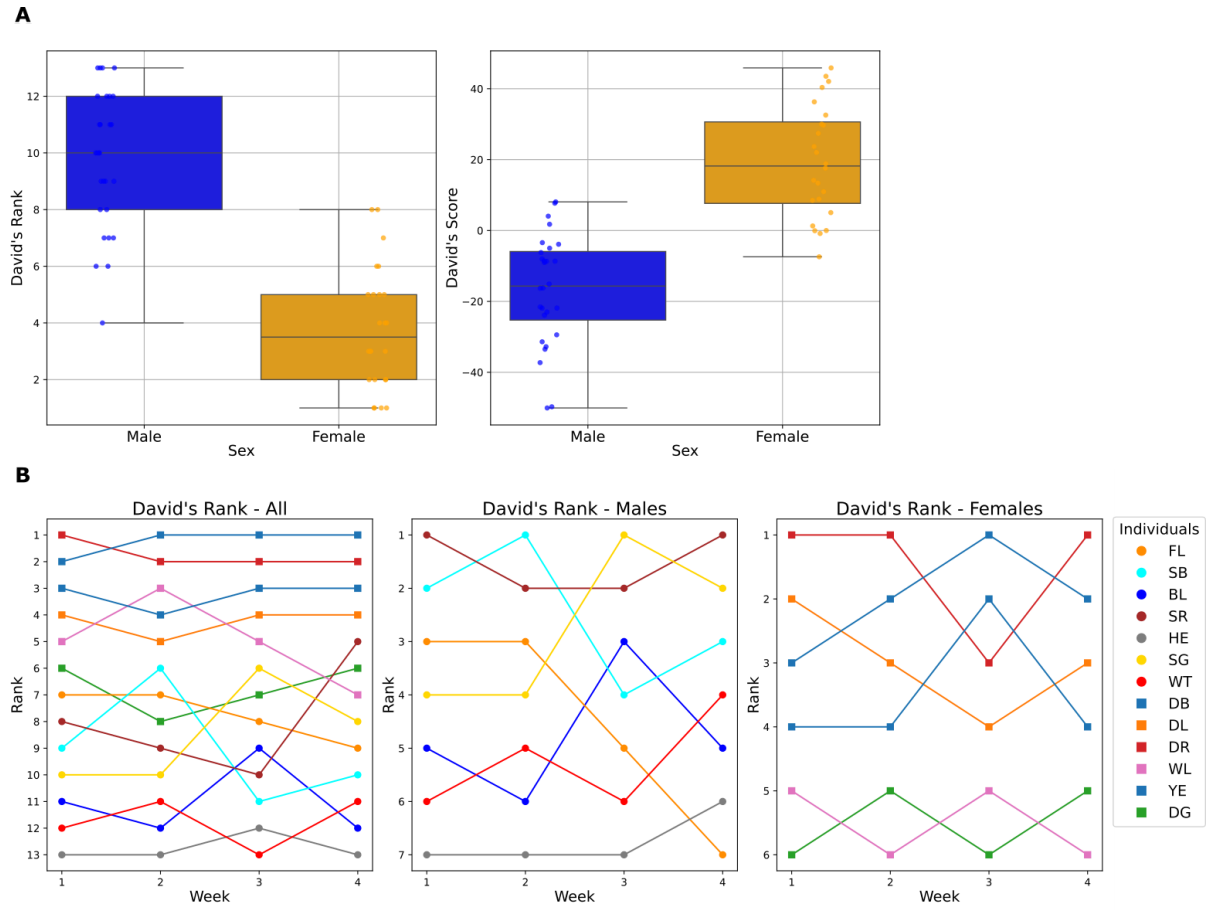


Figure 5: (A) *The combination of scatter and box plots compares male and female dominance scores (or ranks) derived using David's score. The first plot compares the David's ranks, and the second shows the David's scores from which the ranks are derived.* ***(B)*** *These line plots illustrate how each bird's dominance rank changed over time. Each individual has a unique color, with males plotted as circles and females as squares. Separate panels show (i) ranks computed for both sexes combined and (ii) ranks for each sex considered separately.*

Dominance ranks revealed a marked separation between females and males (Figure 5A). Females consistently occupied higher positions in the hierarchy and dominated males in direct interactions (Mann–Whitney U test, $p<0.001$). Over the four-week study, the dominance hierarchy showed minimal fluctuations, with most individuals maintaining stable ranks. Week-to-week comparisons demonstrated that most rank shifts typically involved a gain or loss of just one position (Figure 5B). Repeatability (ICC(1)) values confirmed this stability and were statistically significant across the entire dataset (ICC=0.89, 95% CI [0.77, 0.96], F=33.15, $p<0.001$). Among males,

rank repeatability was moderate but still statistically significant (ICC=0.66, 95% CI [0.31, 0.92], F=8.75, p<0.001), indicating greater rank reshuffling compared to females. Females exhibited higher stability in their order (ICC=0.80, 95% CI [0.51, 0.96], F=17.4, p<0.001), but on the whole, the social structures exhibited broad stability and minimal change in rank over the four-week period.

	Males week 1	Males week 2	Males week 3	Males week 4	Males Combined
Ttri	0.87 (p = 0.11)	0.73 (p = 0.11)	0.74 (p = 0.50)	0.77 (p = 0.36)	0.86 (p = 0.09)
h'	0.68 (p = 0.07)	0.62 (p = 0.11)	0.35 (p = 0.53)	0.41 (p = 0.44)	0.64 (p = 0.11)
steepness	**0.62 (p<0.001)**	**0.57 (p<0.001)**	0.38 (p = 0.12)	0.44 (p = 0.07)	**0.51 (p = 0.03)**
	Females week 1	Females week 2	Females week 3	Females week 4	Females Combined
Ttri	0.75 (p = 0.10)	0.60 (p = 0.89)	0.75 (p = 0.71)	0.60 (p = 0.32)	0.75 (p = 0.67)
h'	0.34 (p = 0.53)	0.31 (p = 0.78)	0.37 (p = 0.54)	0.55 (p = 0.49)	0.43 (p = 0.41)
steepness	0.26 (p = 0.59)	0.22 (p = 0.78)	0.47 (p = 0.07)	0.44 (p = 0.24)	0.43 (p = 0.18)

Table 4: *Value of steepness, h' and Ttri and their respective p-value for observed value for male and female dominance hierarchy. Data are presented separately for each week and combined.*

As predicted for a stable dominance system, the probability of winning an aggressive encounter increased with the rank difference between two individuals (Figure 6a). Although both hierarchies were linear (males: Ttri = 0.86; females: Ttri = 0.75), they were not very steep (males: 0.51; females: 0.43) (Table 4). Thus, even though the male dominance hierarchy was both steeper and more linear overall, both sexes exhibited relatively shallow hierarchies compared to those observed in other species.

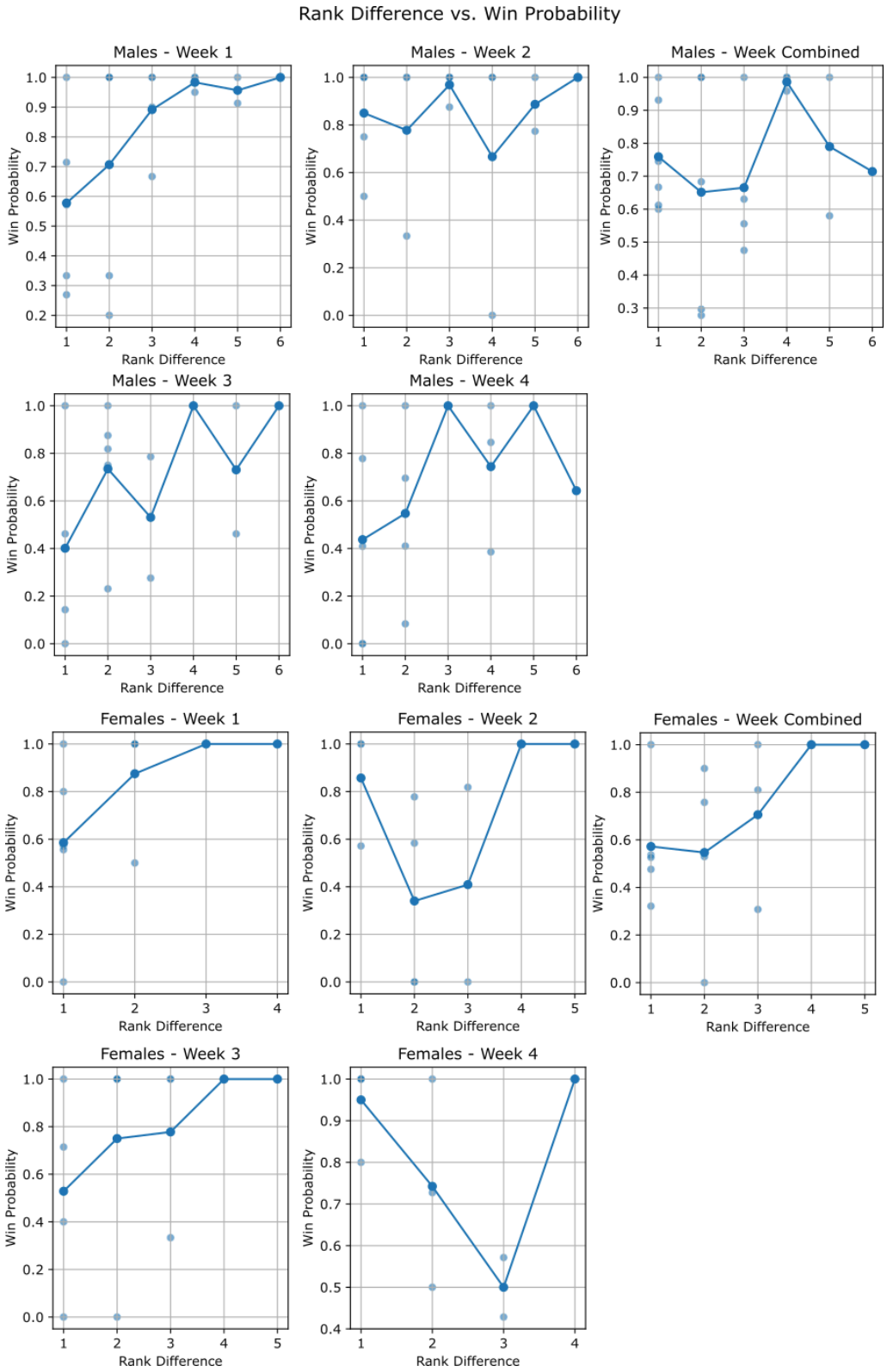


Figure 6 *The scatter plot relates the difference in dominance rank between two birds to the observed probability that the higher-ranked bird wins an encounter. We show separate plots for each week, data combined across weeks, and each sex.*

3.1.6 Displays are negatively correlated with, and may displace aggression

We found a statistically significant negative correlation between display and aggression in the overall dataset (rho = −0.53, p<0.001), as well as in male-male (rho = −0.38, p < 0.001) and female-female (rho = −0.48, p = 0.007) interactions (Table 4; Figure 7). By contrast, all other correlations among the different behaviors were not statistically significant. This strong negative relationship between display and aggression suggested that display may serve as a less costly alternative to aggression, potentially mediating dominance interactions without escalating to physical conflict.

Behavior Pair	Interaction Type	Observed Correlation	P-Value
Display vs. Affiliative	all	0.09	0.37
Display vs. Affiliative	male-male	0.42	0.05
Display vs. Affiliative	female-female	0.23	0.86
Display vs. Affiliative	male-female	0.34	0.27
Display vs. Non-Courtship Aggression	all	**-0.52**	**<0.001**
Display vs. Non-Courtship Aggression	male-male	**-0.37**	**<0.001**
Display vs. Non-Courtship Aggression	female-female	**-0.47**	**0.006**
Display vs. Non-Courtship Aggression	male-female	-0.6	0.49
Affiliative vs. Non-Courtship Aggression	all	-0.58	0.18
Affiliative vs. Non-Courtship Aggression	male-male	0.32	0.14
Affiliative vs. Non-Courtship Aggression	female-female	-0.59	0.17
Affiliative vs. Non-Courtship Aggression	male-female	-0.64	0.63

Table 5 *Correlations between different behavior categories across dyad types.*

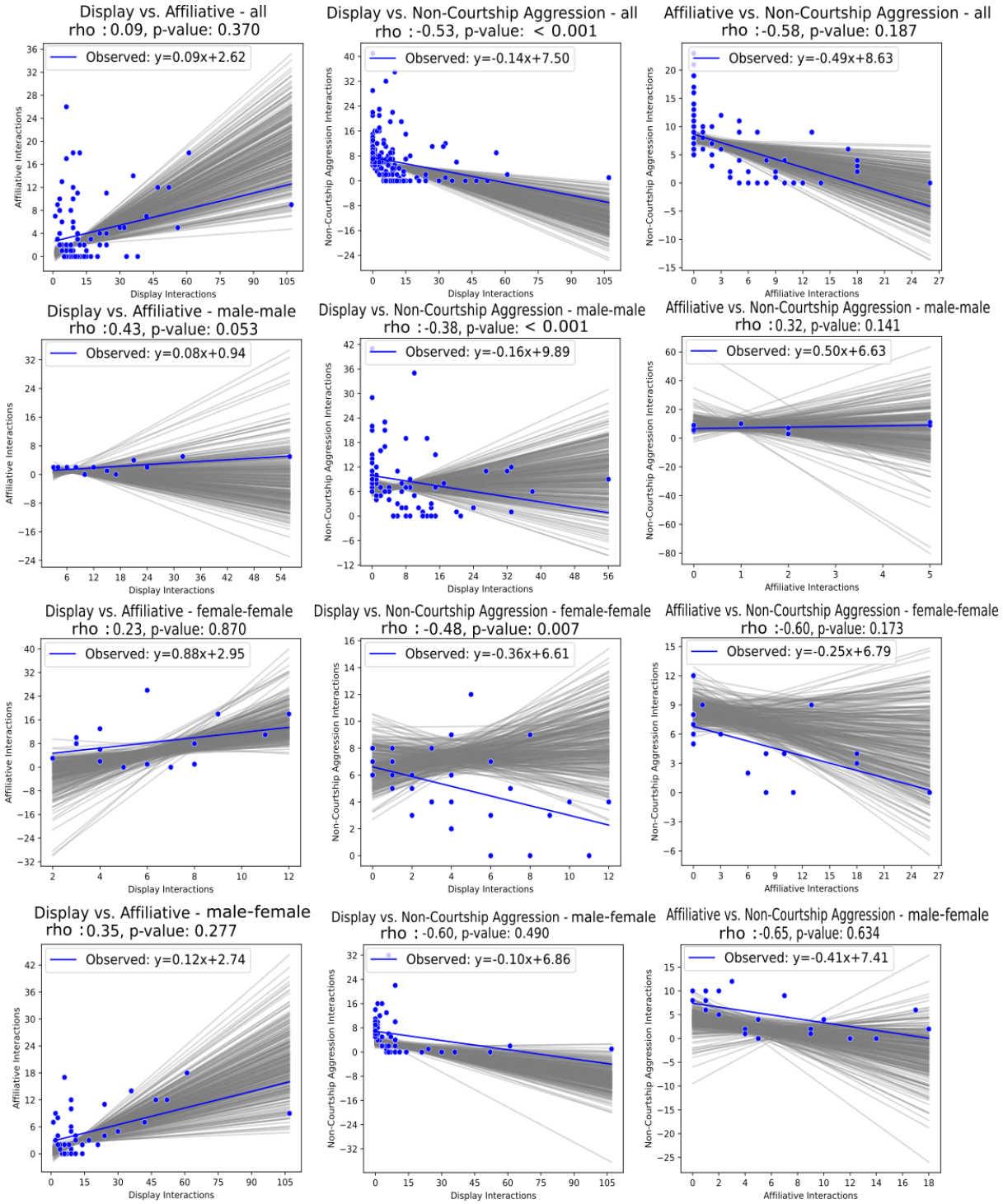


Figure 7: *The scatter plot in blue compares the counts of two different behavior types for the same directional dyad, in a given week. Points represent dyads, the blue line is the best fit to the observed data, and the gray lines are fits to permuted null datasets. Each panel is repeated for a different pair of behavior categories and for the types of dyadic interaction. Plot titles note the observed Spearman correlation (rho) and p value for the observed value of correlation.*

3.1.7 Displays as an appeasement gesture

Finally, we investigated whether the display might function as an appeasement gesture following aggression. Specifically, we measured how often the *recipient* of an aggressive act displayed back to the aggressor and vice versa within fixed time windows of 30 s, 60 s, 90 s, 120 s, and 300 s. Across all time windows, the observed counts were higher than expected under a null model when recipients of aggression performed the display (Figure 8). Thus, the recipient of aggression consistently displayed back to the aggressor more often than chance across time windows and with consistently high effect size, supporting an appeasement-like role for displays to another individual.

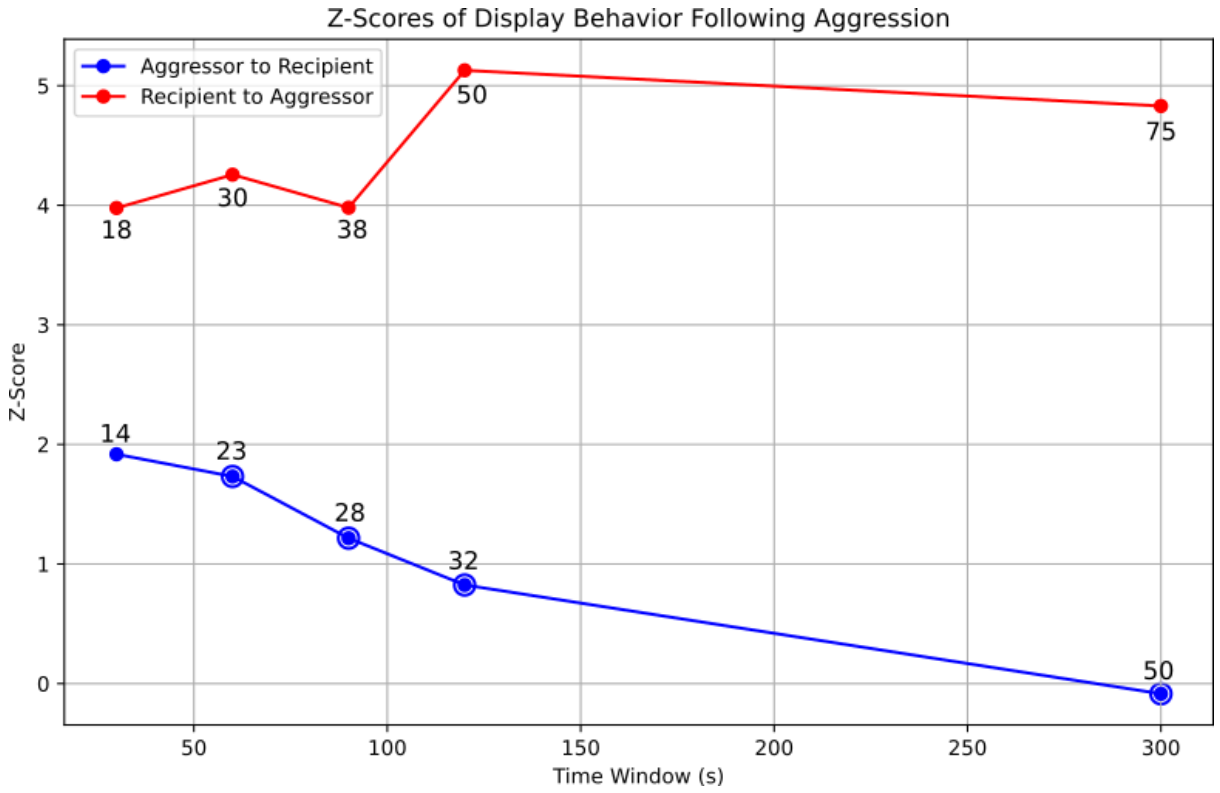


Figure 8: *Z-score of observed display events (by either aggressor or recipient) following an aggressive interaction in a given dyad, across progressively larger time windows after the aggression. The counts of these display events are annotated on the plot, with non-significant Z-scores highlighted by circles.*

However, the absolute numbers of these post-aggression displays remained relatively low, ranging from 18 occurrences at 30 s to 75 occurrences at 300 s across the entire dataset. This suggests that while display reliably occurred after aggression more often than random permutations predict, it was not an overwhelmingly frequent behavior.

3.2 Syntactic analysis of warble song

We analyzed a total of 228 warble sequences for this study, with an average length and standard deviation of 104.86 and 89.24 notes per sequence, respectively. We computed multiple syntax-based distance metrics and then created a pairwise distance matrix for each individual male's weekly repertoire (Supplementary Figure 3a–e).

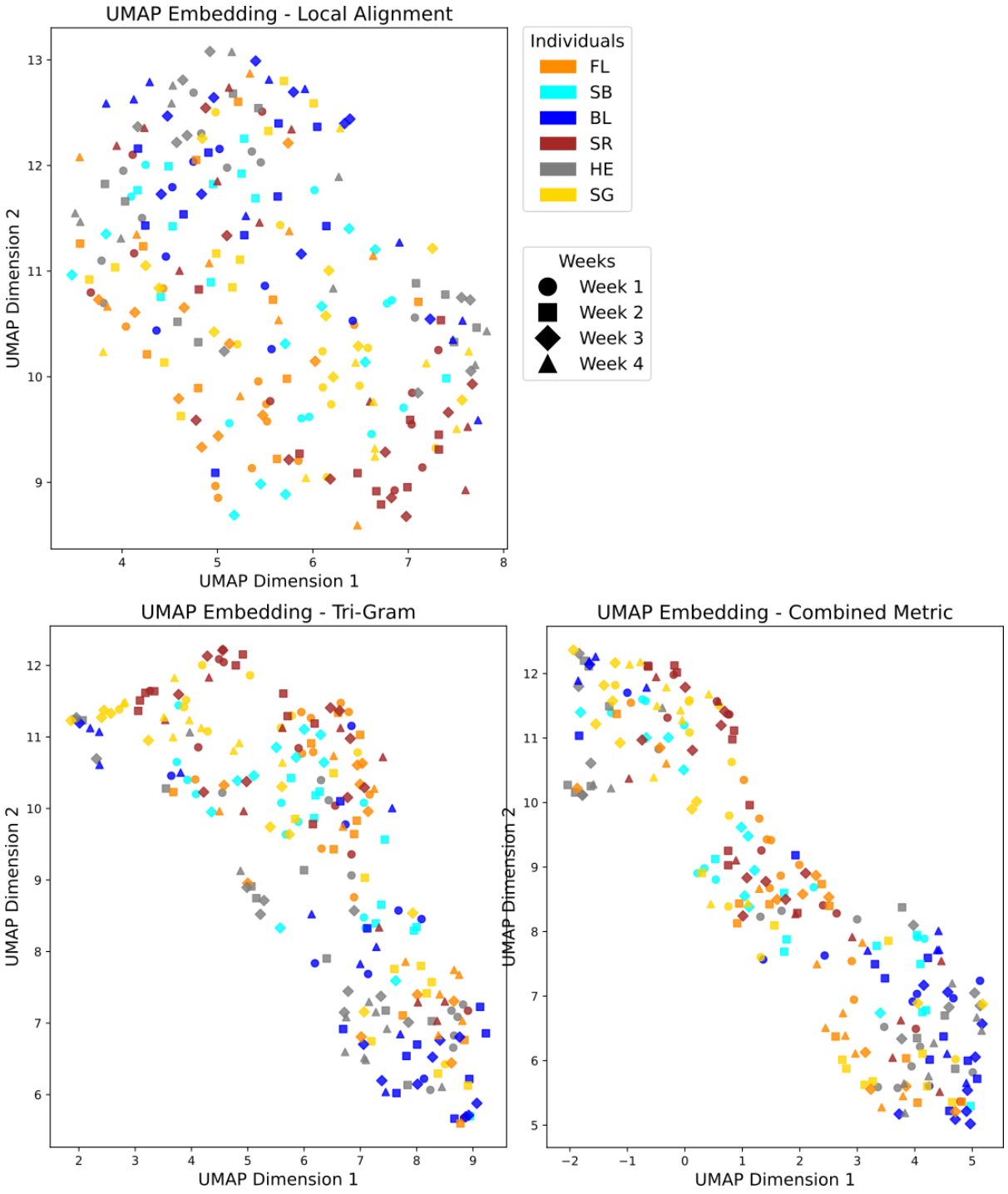


Figure 9: *UMAP projection of a syntax-distance matrix (based on sequence structure) to project each bird's song sequences (pooled across weeks) into two dimensions. Points represent individual sequences, colored by birds and shaped by recording week. Points closer to each other are more similar in syntax.*

3.2.1 Warble syntax exhibits individual signatures, with some temporal drift

To explore individual variation in vocal sequences, we projected the distance matrices from three different syntax metrics onto a two-dimensional UMAP space (Figure 9). This visualization revealed that each bird's warble sequences tended to clump together. Using the combined syntax metric, our statistical tests confirmed a highly significant effect of individual identity on vocal syntax (Figure 10a, z-score: -19.11, $p<0.001$). In other words, warbles produced by the same individual were much more similar to each other than expected by chance. This pattern persisted across the three syntax metrics tested, reinforcing the conclusion that each individual possesses a distinctive syntactic "signature", in addition to colony-level signatures identified previously.

Beyond individual identity, we also detected a weaker yet still significant effect of change across weeks in the combined syntax metric for data across individuals ($p=0.028$, $z=-2.14$; Figure 10b). Although this effect size was substantially smaller than that for the previous test, the results suggest that warble syntax varied across weeks. This trend was consistent across all three syntax metrics, albeit to a lesser degree than individual-specific differences, but consistent with some drift in syntax.

Finally, we assessed whether each bird's warble syntax shifted over time by examining weekly signatures within the repertoire of a single individual (Figure 11; Supplementary Figure 2a–b). Here again, we observed significant effects, consistent again with an individual's syntax "drifting" across the four-week period. However, individual signatures remained strong and their effects were statistically significant.

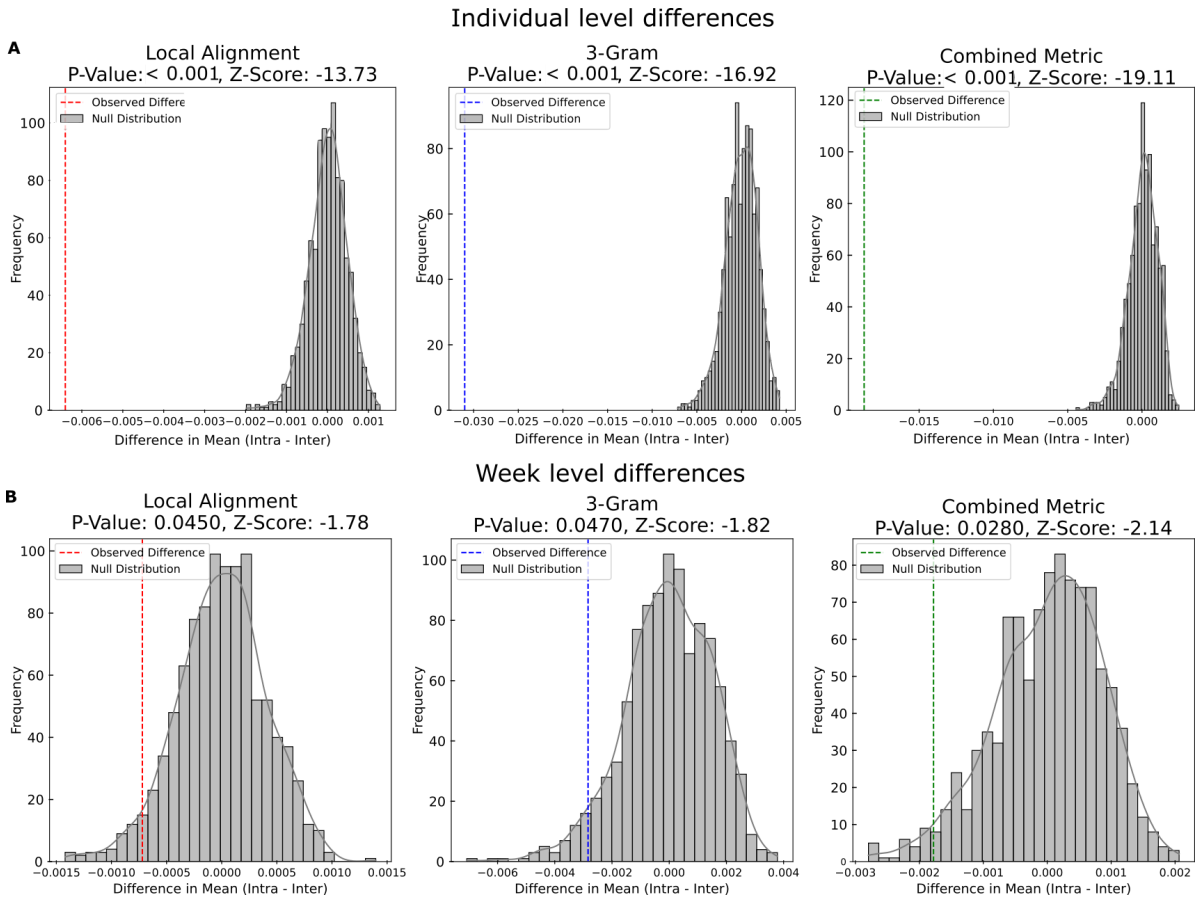


***Figure 10**: Histograms illustrating the difference between mean intra-category distances (e.g. within the same individual or the same week) and mean inter-category distances, compared against null distributions of the same statistic computed from permuted datasets. The observed difference is marked by a dotted line. Titles include the p-value and z-score of the observed value. Separate panels show grouping by individual and by week.*

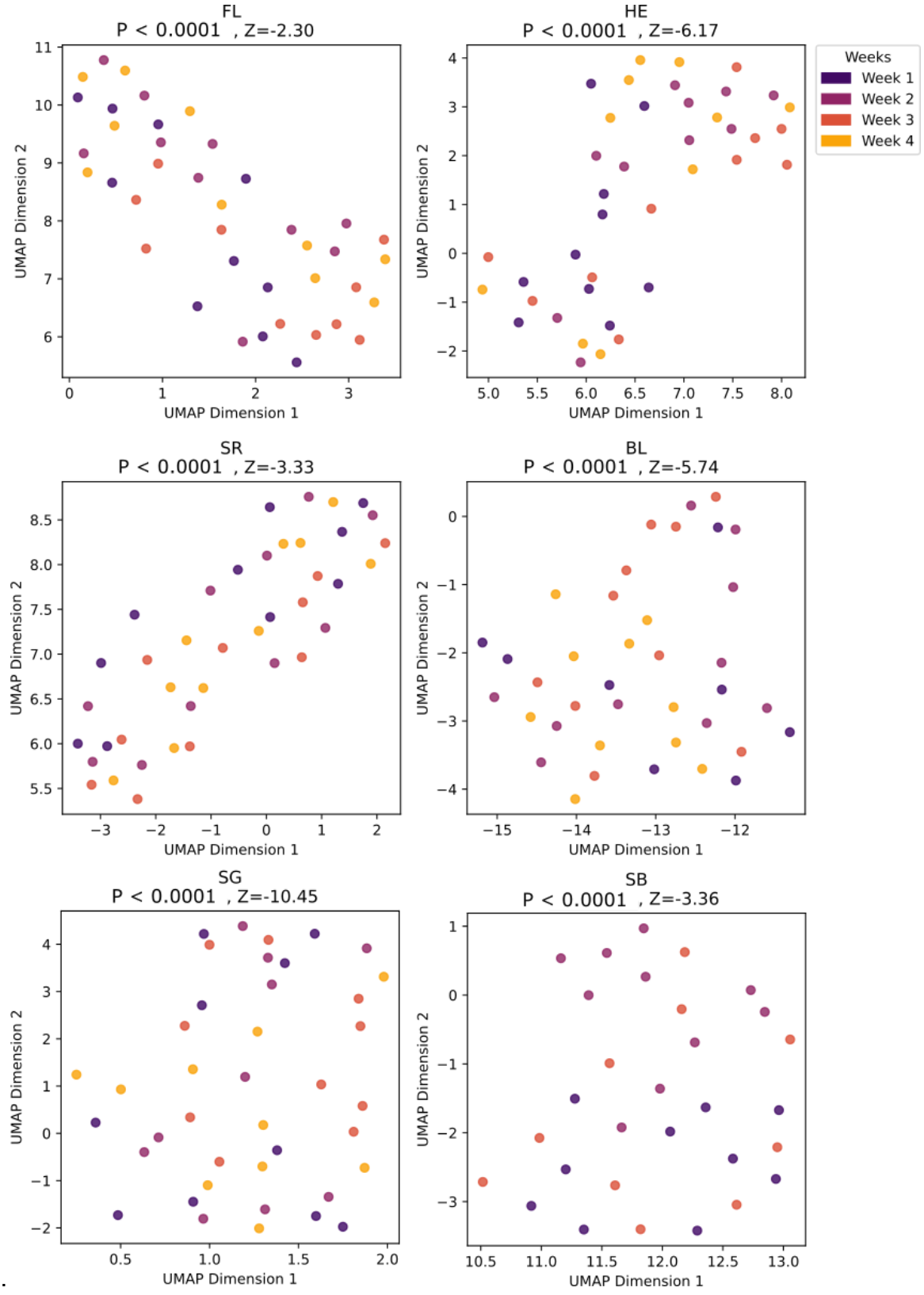
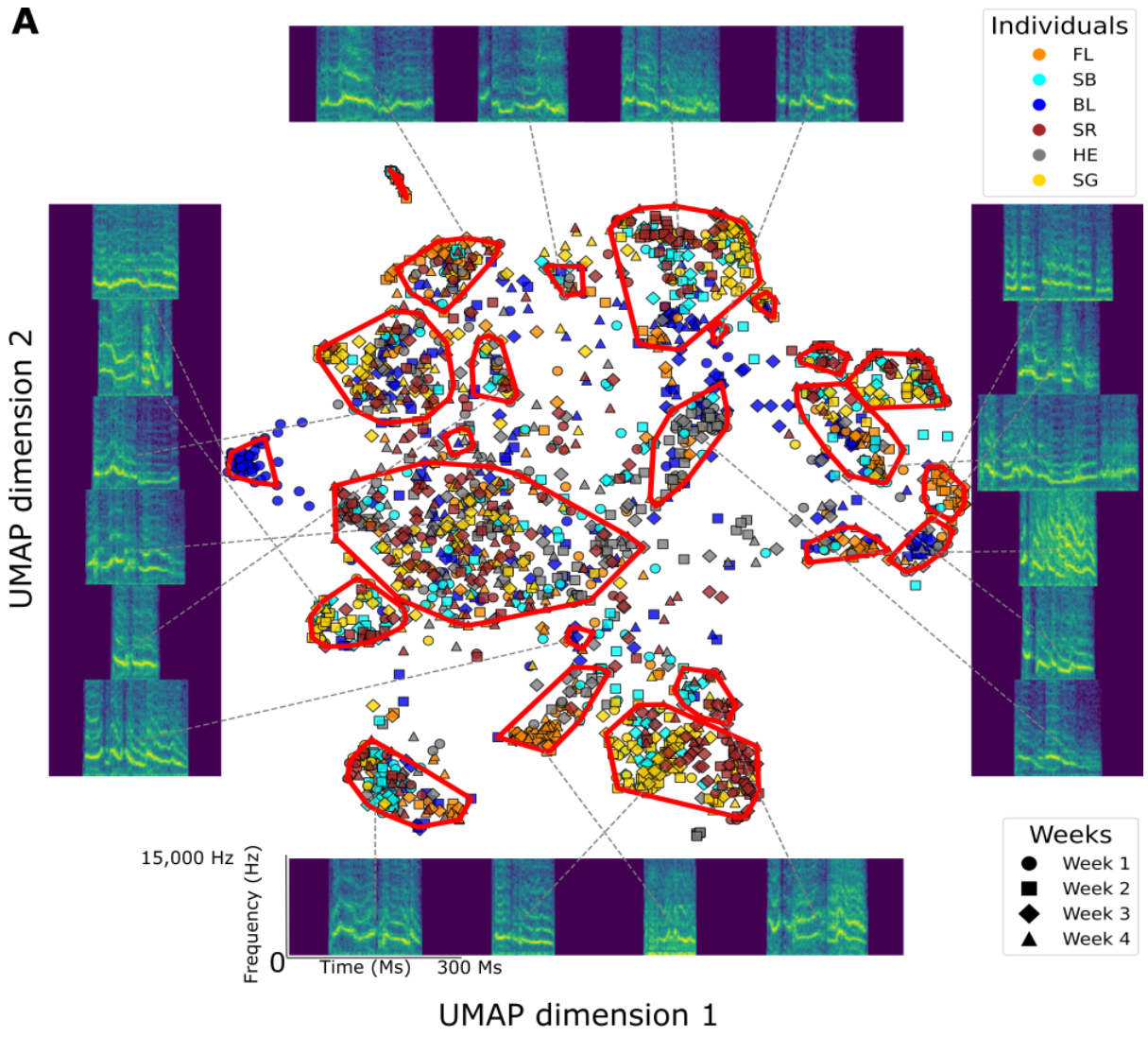


***Figure 11**: UMAP projections of each individual's distance matrix based on the combined syntax metric, Each point is colored by week, illustrating how an individual's syntax changes over time.*

3.3 Spectral analysis of contact call-like elements



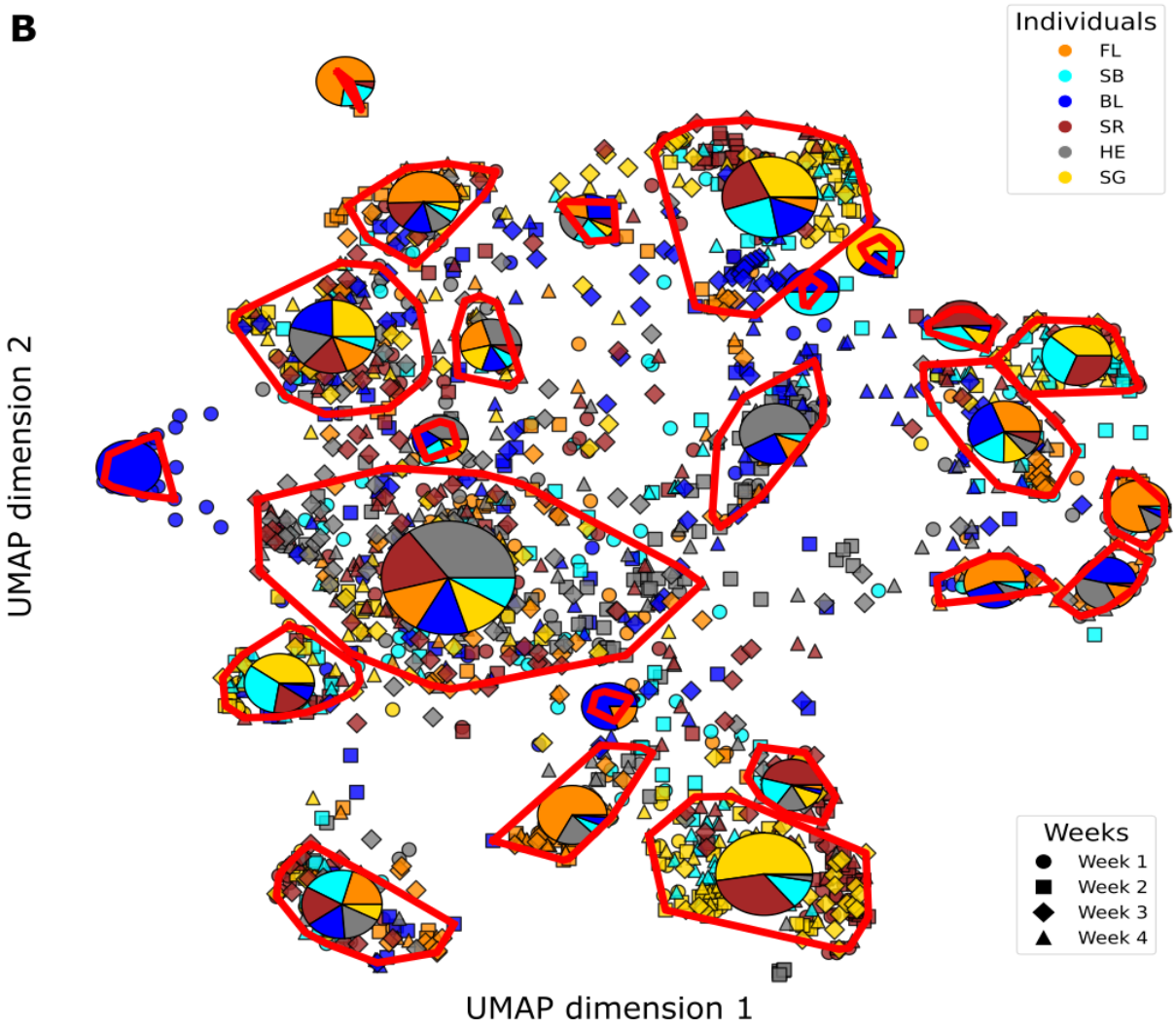


***Figure 12**: 2D UMAP projection of spectral similarity using DTW-based distances on a sample of 100 calls per bird. Each point is a single call, colors represent an individual bird and shapes represent week of data. Red lines outline convex hulls for visualizing clusters. (A) sample spectrograms illustrate how calls differ among clusters. (B) Pie charts indicate the proportion of each cluster occupied by an individual, placed atop the UMAP clusters and scaled by the number of calls in each cluster.*

From the 228 warble sequences used for syntax analysis, we identified 5827 b notes (contact call like elements) for further investigation of changes in time-frequency structure, representing 24% of all notes in our warble dataset. These b notes were then examined to assess individual and week-level acoustic signatures.

3.3.1 Contact call repertoire exhibits individual signatures, with temporal drift across weeks

We first assessed whether all b notes clustered according to individual identity (Figure 13a). The effect was statistically significant ($p<0.001$, $z=-19.4$), indicating that calls from a given bird tended to be more spectrally similar to each other than to calls from other birds. When we focused on calls within each cluster—groupings of spectrally similar b notes (or subtypes) —individual signatures were statistically significant at this level as well ($p<0.001$, $z=-24.2$).

Next, we evaluated week-level variation in structures of b notes, to examine whether there was also spectral drift in addition to changes in syntax (Figure 13b). A combined analysis for all b notes uncovered a significant effect ($p<0.001$, $z=-4.2$), implying that b notes shift slightly from week to week. Within clusters, however, there was no significant change across weeks ($p=0.512$, $z=0.01$), suggesting that the week-level separation in the full dataset might reflect changing proportions of different call subtypes rather than acoustic changes within individual subtypes. Overall, these findings mirror our syntax analysis results, suggesting strong individual acoustic signatures, retained across weeks even in the face of some drift in sequence structure and call use. We further explored whether each bird's acoustic features changed over time by systematically examining the properties of b notes week by week (Figure 14). The UMAP visualization suggested that the effect we observed above was, indeed, because individual birds sometimes shifted their usage of particular call subtype clusters across weeks.

Based on the findings above, we next investigated whether certain individuals shared particular call subtypes. Heat maps of raw call counts (Figure 15) showed that some birds exhibited greater overlap in their cluster usage than others. This pattern is also evident in the repertoire similarity network (Figure 17; Supplementary Figure 3f), where edges connect individuals who share more b note clusters. Together, these results indicate that certain pairs of birds converge on similar call subtypes or at least use overlapping repertoires more frequently than expected by chance. This in turn motivated our next set of analyses.

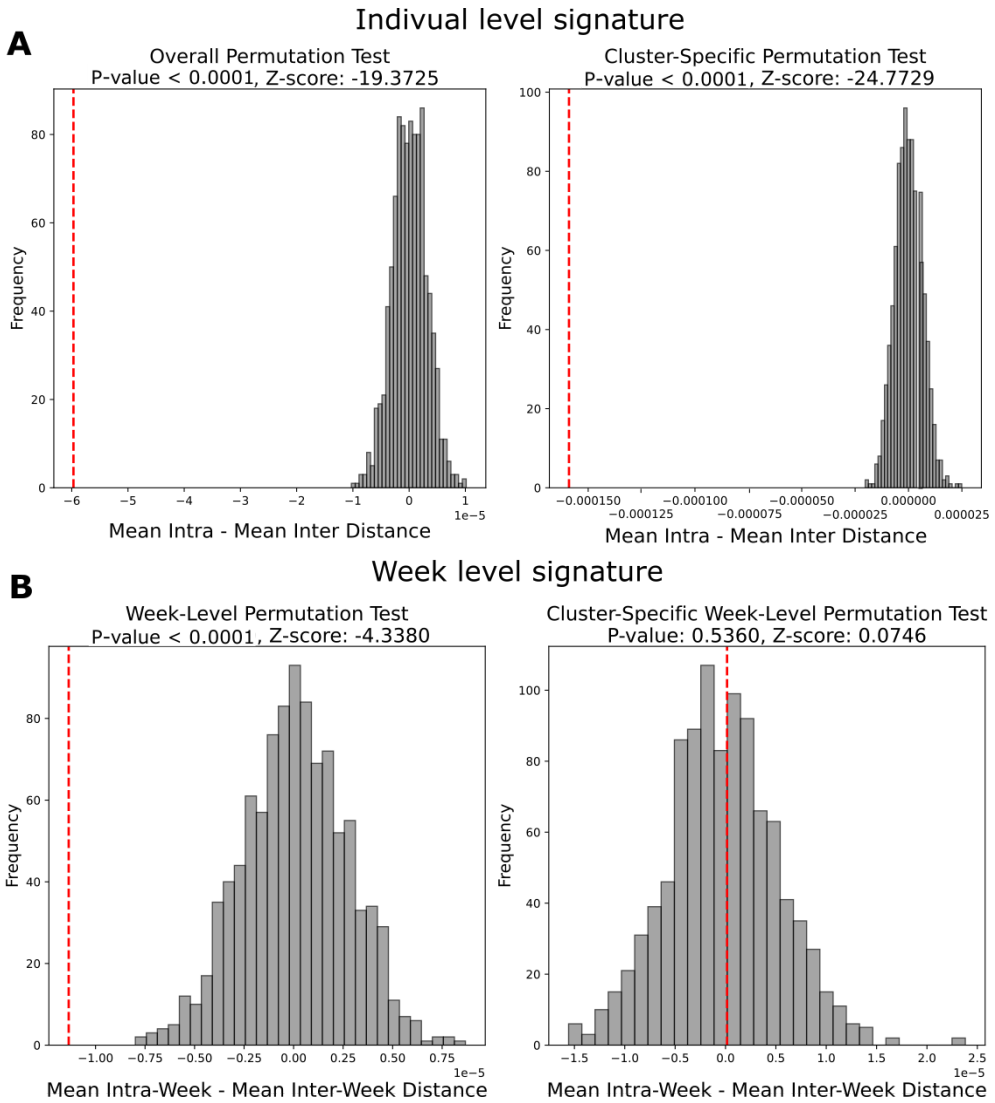


***Figure 13**: Histogram showing the difference between intra- and inter-category mean distances (either at the (A) individual and (B) week level), compared with null distributions derived from permutations. The observed difference is a dotted red line, and titles show p-values and z-scores. Analyses are repeated for overall repertoire distances and distances within each cluster.*

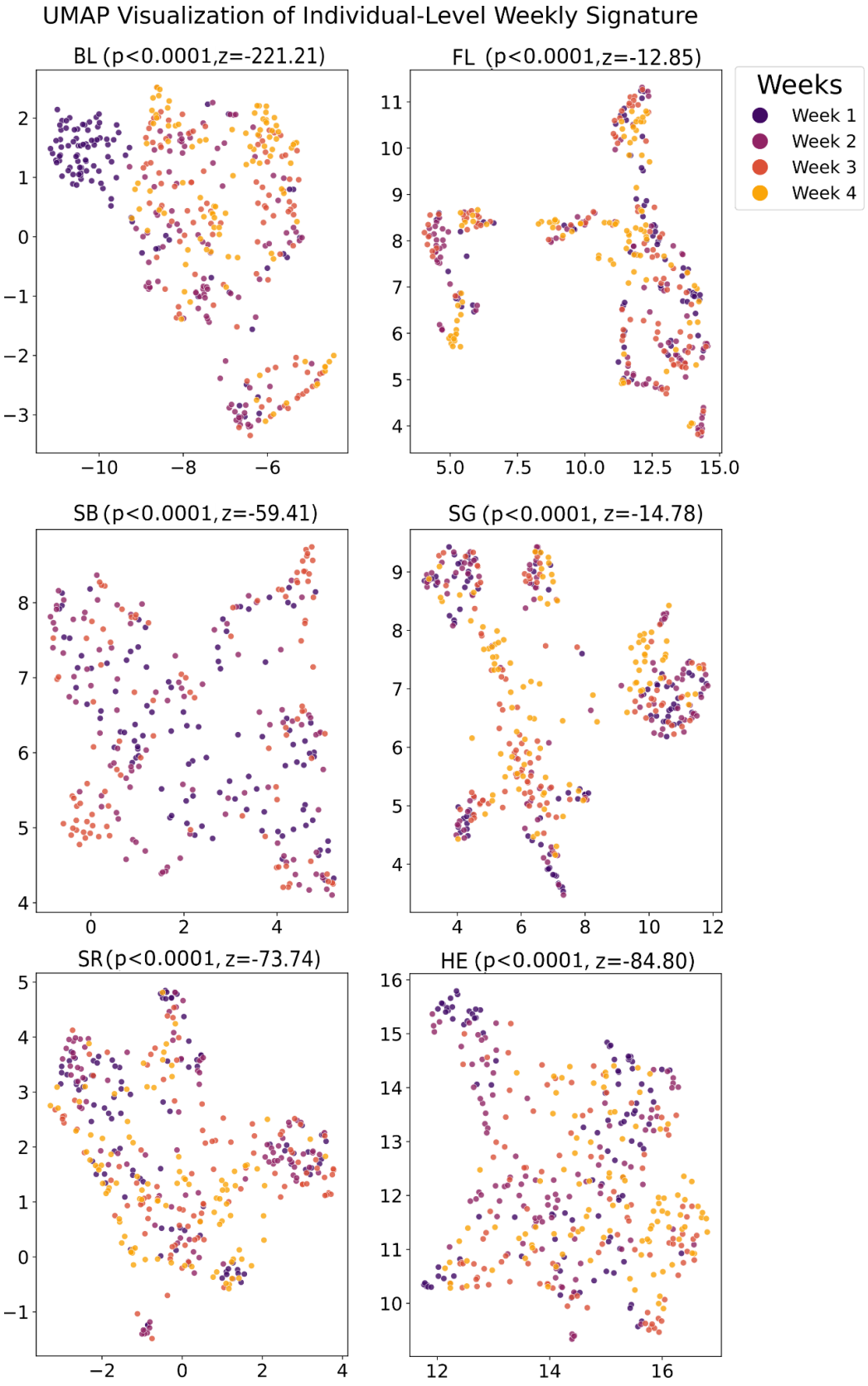


***Figure `14**: 2D UMAP projection of each bird's calls across weeks.*

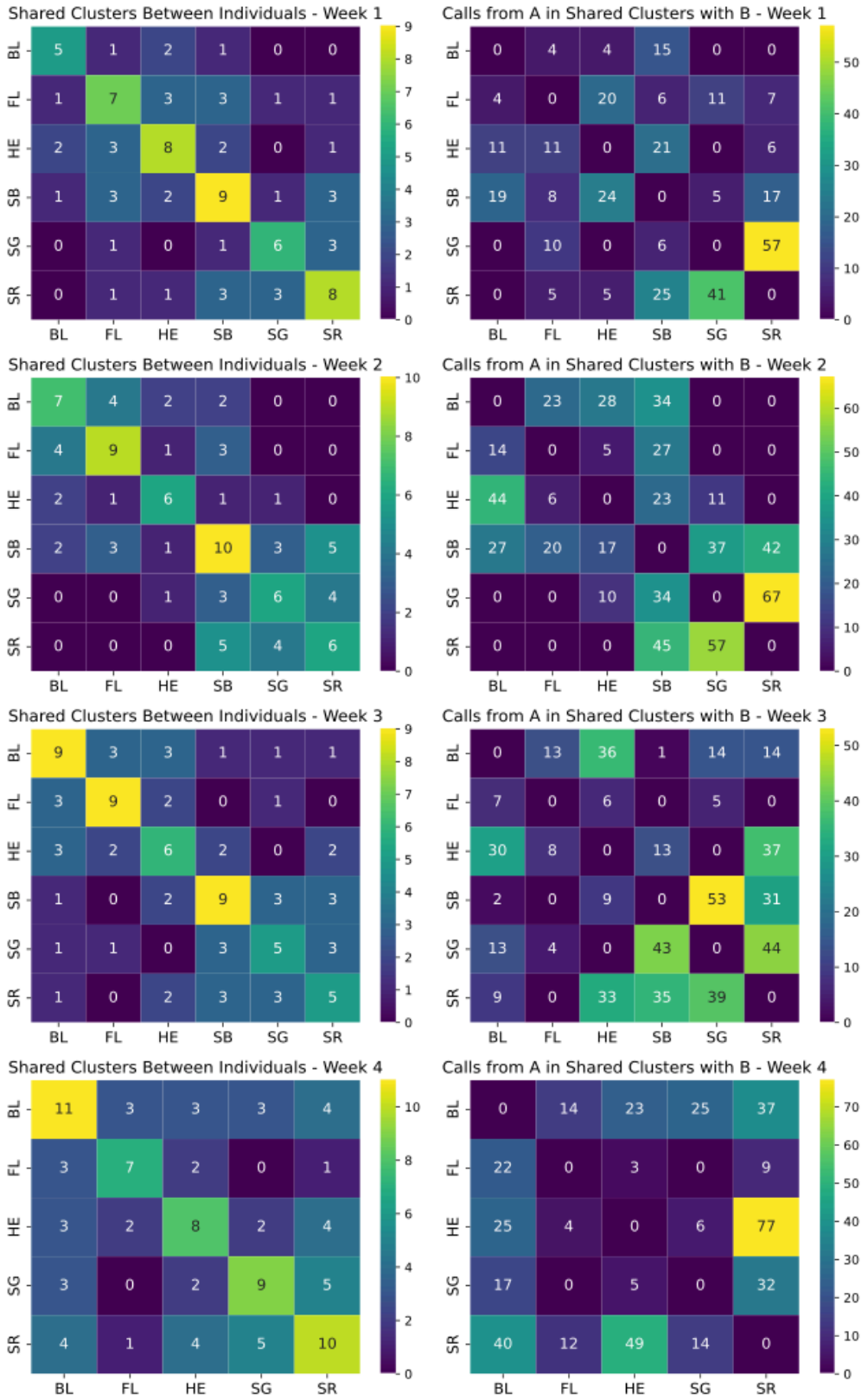


Figure 15: *Heat map showing call clusters and the number of calls within those clusters shared across different weeks between all pairwise combinations of individuals.*

3.4 Social network effects on acoustic signals

To recap our methods, we next examined relationships between the following set of metrics:

Acoustic metrics

1. Syntax metrics:
    - Four independent syntax metrics were derived from warble sequences, and one combined metric.
2. Spectral structure metrics:
    - Two measures of overall spectral properties of each bird's repertoire, across and within clusters.
3. Call subtype use metrics:
    - Two measures quantifying how often birds shared particular subtypes of contact calls.

Interactions measured from the social network

We next explored how these acoustic features related to social behavior at the pairwise (dyadic) level, focusing on four main variables:

1. Number of display interactions (Figure 16a)
2. Number of displays to the same female (Figure 16b)
3. Number of aggressive interactions (Figure 16c)
4. Dominance rank difference (Figure 16d)

3.4.1 Relationships between acoustic features and edge parameters of the social network

1. Number of display interactions: males that display to each other more diverge in contact call subtype use (Figure 17a).

- **Syntax**: The correlation coefficient showed inconsistent results, with effects in both directions and inconsistent significance across different metrics: local alignment (rho = -0.12, p = 0.18), co-occurrence ratio (rho = 0.04, p = 0.45), co-occurrence count (rho = 0.10, p = 0.45), tri-gram (rho = 0.17, p = 0.009), and combined metric (rho = -0.01, p = 0.02). Thus, we did not detect a consistent effect of male-male display on syntax similarity.
- **Spectral structure**: Although both tests showed significant results for both metrics, they indicated effects in opposite directions. For slope, the first test was significant for within-cluster distance (z-score = -1.58, p = 0.04) and overall distance (z-score = -3.08, p < 0.001), indicating that the observed value is less than chance. However, for the correlation coefficient, the results indicated a value greater than chance, both for within-cluster distance (z-score = 4.24, p < 0.001) and overall distance (z-score = 2.47, p = 0.003). Thus, as the two tests indicated effects in opposite directions, we considered this to be insufficient evidence of a consistent effect of male-male display on the spectral structure of contact calls.
- **Call subtype Use**: Both metrics exhibited a negative relationship for both tests used. For the number of shared calls in a cluster, both tests showed a significant and negative relationship: slope (z-score = -2.12, p = 0.017) and correlation coefficient (rho = -0.25, z-score = -4.48, p < 0.001). For the number of shared clusters, although both tests indicated an effect in the same direction, the slope was not significant: slope (z-score = -1.15, p = 0.11) and correlation coefficient (rho = -0.14, z-score = -2.92, p = 0.002). As both tests indicate the same direction of effect and all tests were significant except the slope test for one of the two metrics, we considered this evidence of a consistent negative effect. Therefore, males engaging in increased mutual displays diverged in their use of call subtypes.

2. Number of displays to the same females: males displaying more to the same females diverge syntactically, but converge in their use of contact-call subtypes (Figure 17b).

- **Syntax**: All five syntax metrics exhibited a significant positive relationship across both tests. For the correlation coefficient test, test results were as follows: local alignment (rho = 0.09, z-score = 3.39, p = 0.001), co-occurrence count (rho = 0.17, z-score = 7.25, p < 0.001), co-occurrence ratio (rho = 0.12, z-score = 8.83, p < 0.001), tri-gram (rho = 0.16, z-score = 8.03, p < 0.001), and combined metric (rho = 0.27, z-score = 10.47, p < 0.001). All tests for slope also indicated a significant positive effect. This implied that males displaying to the same female were divergent in the syntactic structure of their warble sequences

- **Spectral structure**: We found no significant effects, as the observed effects were more likely than chance but were in a negative direction. Thus, the effect size and observed value had opposite signs. For the correlation coefficient metric, results were as follows: within-cluster distance (rho = -0.07, z-score = 4.94, p < 0.001). For slope: within-cluster distance (slope = -9.43E-07, z-score = 4.8, p < 0.001) and overall distance (slope = -1.11E-06, z-score = 3.43, p < 0.001). We interpreted this as inconsistent evidence of an effect.

- **Call subtype use**: Both of our tests indicated a positive effect for both metrics used. Although the correlation coefficient was significant for both metrics—number of shared calls in clusters (rho = 0.33, z = 3.92, p < 0.001) and number of clusters (rho = 0.20, z = 2.72, p = 0.005)—the results from the slope test, although indicating an effect in the same direction, were non-significant: both for number of shared calls in clusters (z = 0.55, p = 0.28) and number of clusters (z = 0.15, p = 0.44). As both tests indicated results in the same direction, and the correlation coefficient test was significant, according to our criteria, we interpreted this as a consistent positive effect. Therefore, males that displayed to the same female(s) more frequently tended to converge in call type use.

.

3. Number of aggressive interactions: Aggressive interactions do not influence the vocal repertoire (Figure 17c)

- **Syntax**: Tests for the correlation coefficient were inconsistent across different metrics in both direction and significance: local alignment ($z = 1.8$, $p = 0.03$), co-occurrence count ($z = -2.48$, $p = 0.007$), co-occurrence ratio ($z = -0.28$, $p = 0.39$), trigram ($z = -4.23$, $p < 0.001$), and combined metric ($z = -1.57$, $p = 0.06$).

- **Spectral:** Although both tests across the metrics recovered a negative relationship, significance was found for only one metric based on the correlation coefficient: within-cluster distance ($z = -2.94$, $p = 0.002$), but not for the other metric, overall distance ($z = -0.52$, $p = 0.3$). Conservatively, we interpreted this as inconsistent evidence of an effect.

- **Call subtype use**: Both metrics were significantly positively correlated with the number of aggressive interactions, but the effect shown by the slope was in the opposite direction. For the number of shared calls in a cluster, the slope test ($z = -1.46$, $p = 0.08$) suggested that the observed effect was less likely than chance, whereas the correlation test returned a significant positive value ($rho = 0.2$, $z = 1.97$, $p = 0.025$). Thus, as both tests indicated opposite directions for the effect, we interpreted this as the lack of a consistent relationship between aggression and call subtype use.

4. Dominance rank difference: individuals closer in the hierarchy exhibit weak evidence of syntactic similarity (Figure 17d).

- **Syntax**: For four out of five metrics, we observed a weak significant or marginally insignificant positive relationship across both tests—positive for co-occurrence ratio (slope: $z = 1.59$, $p = 0.07$; rho: $z = 1.82$, $p = 0.05$), co-occurrence count (slope: $z = 1.41$, $p = 0.088$; rho: $z = 1.50$, $p = 0.073$), tri-gram distance (slope: $z = 1.45$, $p = 0.086$; rho: $z = 1.95$, $p = 0.03$), and combined metric (slope: $z = 1.61$, $p = 0.062$; rho: $z = 1.89$, $p = 0.036$). The relationship was not significant for local alignment (slope: $z = -0.67$, $p = 0.25$; rho: $z = -1.14$, $p = 0.13$). Thus, as four out of five syntax metrics showed a positive relationship with dominance rank difference, with all tests close to the significance threshold, we interpreted this as weak evidence for a positive relationship between rank difference and distance in syntactic structure.

- **Spectral:** Inconsistent results were observed in terms of significance for the correlation coefficient test across the two metrics used: within-cluster distance ($p = 0.02$) and overall distance ($p = 0.26$).

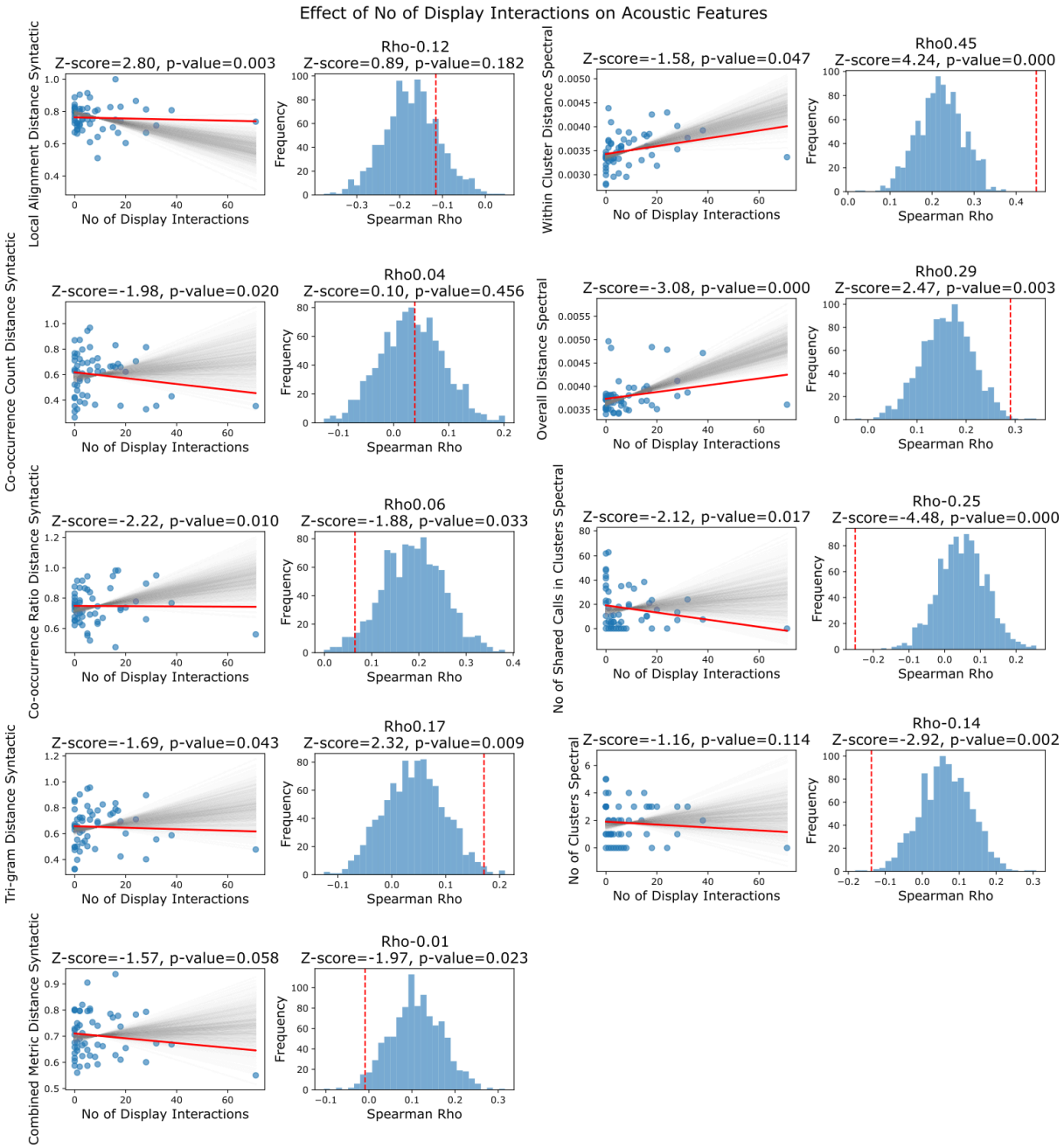
- **Call subtype**The correlation test returned non-significant or marginally insignificant results across both metrics used: number of shared calls in a cluster ($p = 0.16$) and number of clusters ($p = 0.052$).

Overall, rank difference exerted weak effects only on warble syntax: individuals of similar rank displayed greater syntactic similarity, but the effect size was low.

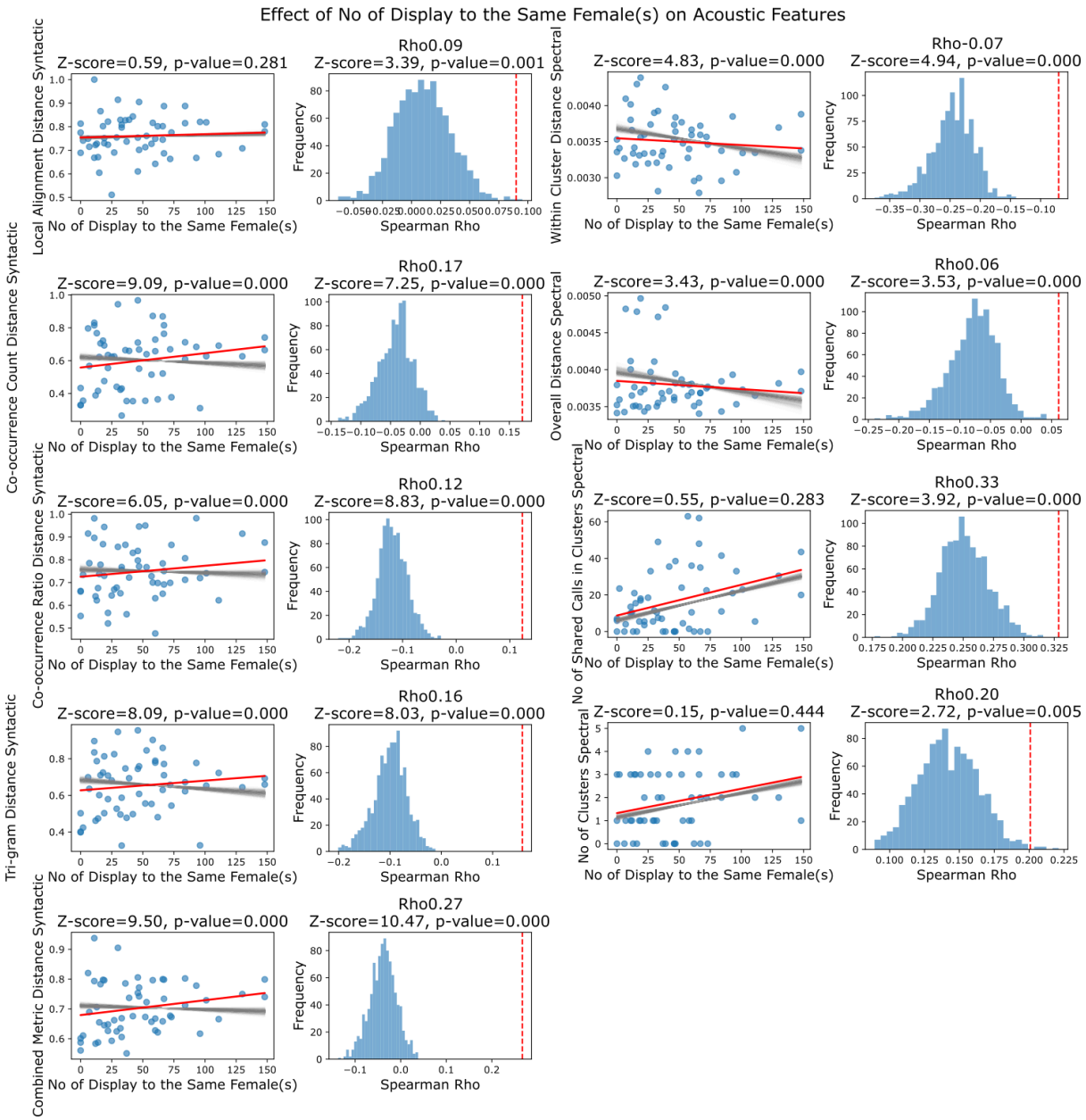
Summary of dyad-level findings:

1. Male-male display interactions were correlated to divergence in call subtype use.
2. Males that displayed to the same female(s) were convergent in call subtype usage but divergent in syntax.
3. We found weak evidence that birds with closer dominance ranks exhibited higher syntactic similarity.

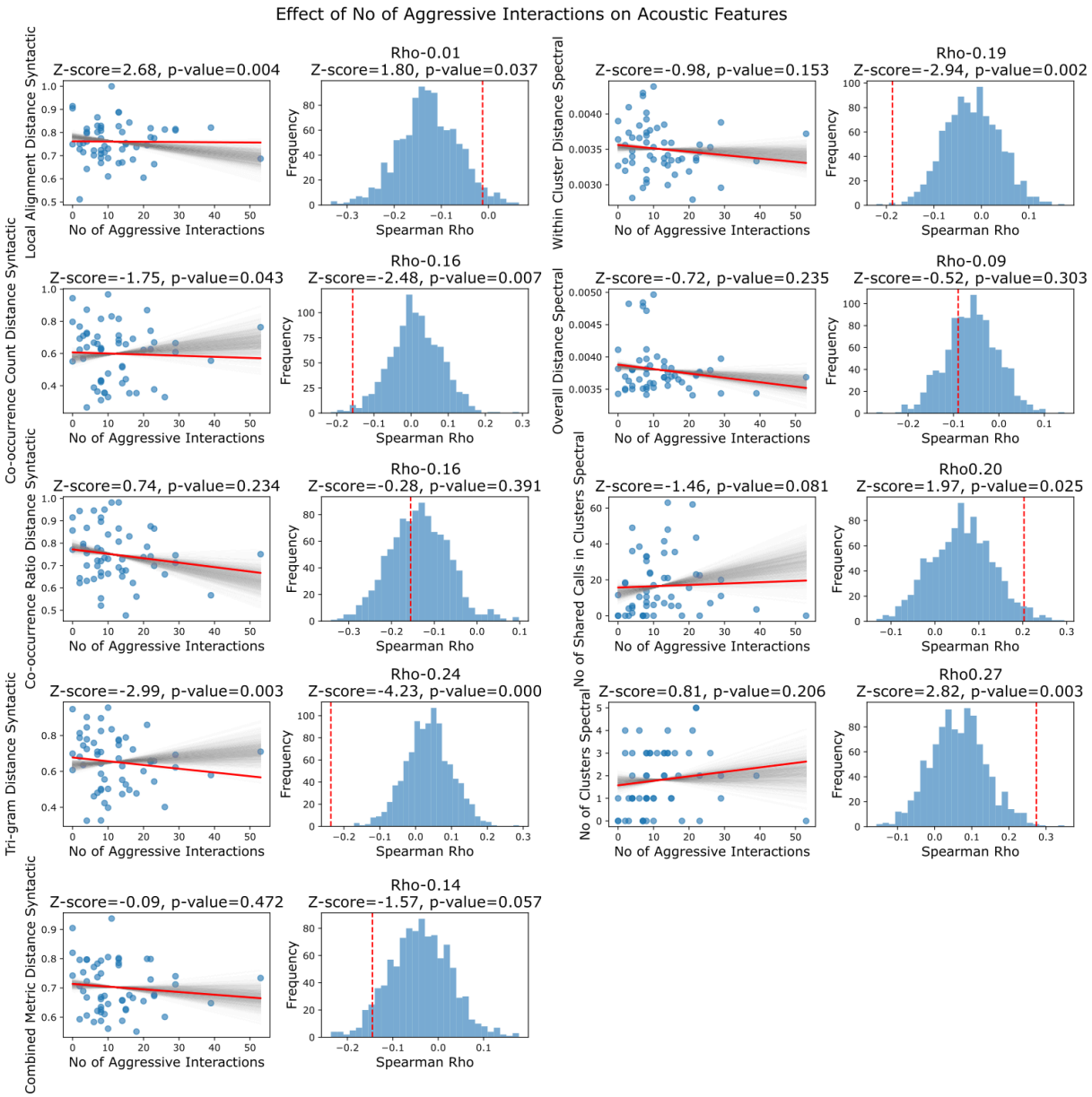
A


Effect of No of Display Interactions on Acoustic Features
Local Alignment Distance Syntactic
Z-score=2.80, p-value=0.003
Rho-0.12
Z-score=0.89, p-value=0.182
Within Cluster Distance Spectral
Z-score=-1.58, p-value=0.047
Rho0.45
Z-score=4.24, p-value=0.000
Co-occurrence Count Distance Syntactic
Z-score=-1.98, p-value=0.020
Rho0.04
Z-score=0.10, p-value=0.456
Overall Distance Spectral
Z-score=-3.08, p-value=0.000
Rho0.29
Z-score=2.47, p-value=0.003
Co-occurrence Ratio Distance Syntactic
Z-score=-2.22, p-value=0.010
Rho0.06
Z-score=-1.88, p-value=0.033
No of Shared Calls in Clusters Spectral
Z-score=-2.12, p-value=0.017
Rho-0.25
Z-score=-4.48, p-value=0.000
Tri-gram Distance Syntactic
Z-score=-1.69, p-value=0.043
Rho0.17
Z-score=2.32, p-value=0.009
No of Clusters Spectral
Z-score=-1.16, p-value=0.114
Rho-0.14
Z-score=-2.92, p-value=0.002
Combined Metric Distance Syntactic
Z-score=-1.57, p-value=0.058
Rho-0.01
Z-score=-1.97, p-value=0.023
No of Display Interactions
Spearman Rho
Frequency

B


Effect of No of Display to the Same Female(s) on Acoustic Features
Local Alignment Distance Syntactic
Z-score=0.59, p-value=0.281
Rho0.09
Z-score=3.39, p-value=0.001
Within Cluster Distance Spectral
Z-score=4.83, p-value=0.000
Rho-0.07
Z-score=4.94, p-value=0.000
Co-occurrence Count Distance Syntactic
Z-score=9.09, p-value=0.000
Rho0.17
Z-score=7.25, p-value=0.000
Overall Distance Spectral
Z-score=3.43, p-value=0.000
Rho0.06
Z-score=3.53, p-value=0.000
Co-occurrence Ratio Distance Syntactic
Z-score=6.05, p-value=0.000
Rho0.12
Z-score=8.83, p-value=0.000
No of Shared Calls in Clusters Spectral
Z-score=0.55, p-value=0.283
Rho0.33
Z-score=3.92, p-value=0.000
Tri-gram Distance Syntactic
Z-score=8.09, p-value=0.000
Rho0.16
Z-score=8.03, p-value=0.000
No of Clusters Spectral
Z-score=0.15, p-value=0.444
Rho0.20
Z-score=2.72, p-value=0.005
Combined Metric Distance Syntactic
Z-score=9.50, p-value=0.000
Rho0.27
Z-score=10.47, p-value=0.000
No of Display to the Same Female(s)
Frequency
Spearman Rho

C


Effect of No of Aggressive Interactions on Acoustic Features
Local Alignment Distance Syntactic
Z-score=2.68, p-value=0.004
Rho-0.01
Z-score=1.80, p-value=0.037
Within Cluster Distance Spectral
Z-score=-0.98, p-value=0.153
Rho-0.19
Z-score=-2.94, p-value=0.002
Co-occurrence Count Distance Syntactic
Z-score=-1.75, p-value=0.043
Rho-0.16
Z-score=-2.48, p-value=0.007
Overall Distance Spectral
Z-score=-0.72, p-value=0.235
Rho-0.09
Z-score=-0.52, p-value=0.303
Co-occurrence Ratio Distance Syntactic
Z-score=0.74, p-value=0.234
Rho-0.16
Z-score=-0.28, p-value=0.391
No of Shared Calls in Clusters Spectral
Z-score=-1.46, p-value=0.081
Rho0.20
Z-score=1.97, p-value=0.025
Tri-gram Distance Syntactic
Z-score=-2.99, p-value=0.003
Rho-0.24
Z-score=-4.23, p-value=0.000
No of Clusters Spectral
Z-score=0.81, p-value=0.206
Rho0.27
Z-score=2.82, p-value=0.003
Combined Metric Distance Syntactic
Z-score=-0.09, p-value=0.472
Rho-0.14
Z-score=-1.57, p-value=0.057
No of Aggressive Interactions
Frequency
Spearman Rho

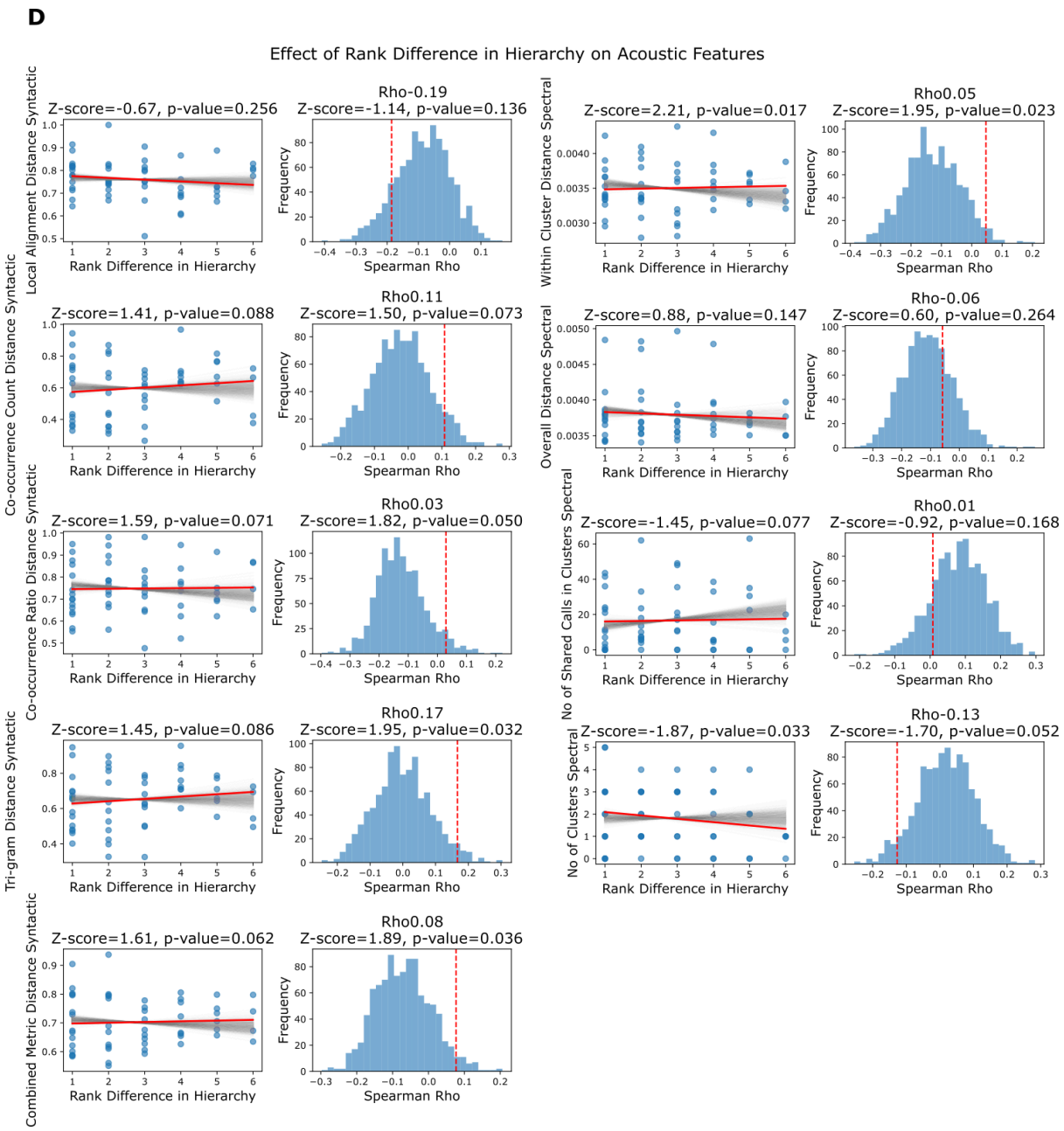
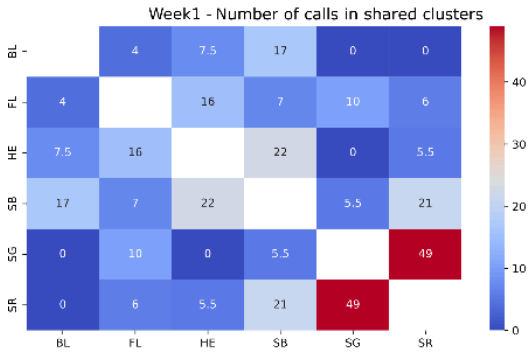


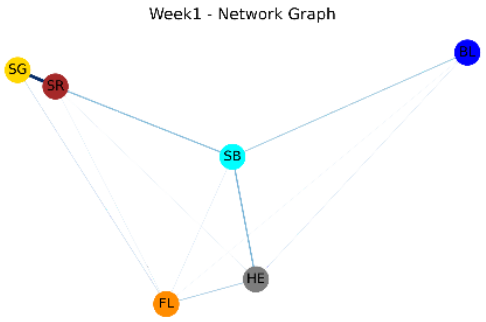
Figure 16: *Comparison of an acoustic similarity or distance measure to a selected behavioral measure (e.g., aggression, display). Two types of comparisons between the two variables are shown. First, a simple regression, with results plotted as a line of best fit on the scatter plot of observed data, and the fit to permuted datasets shown in light grey. The Z-score and p-value of the observed slope are reported in the plot title. The second comparison uses Spearman's rank correlation coefficient (rho), where the observed correlation coefficient is plotted as a line on the histogram of correlation coefficients obtained from permuted datasets. The value of the observed rho, its p-value, and Z-score are reported in the plot title. Separate sub-figures show the effects of different behavioral measures on acoustic features: (A) Display, (B) Display to the same female(s), (C) Aggression, and (D) Rank Difference.*

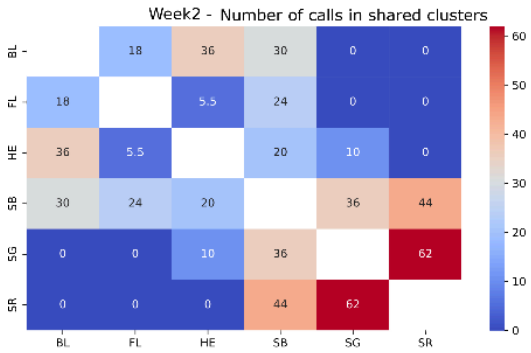
3.4.3 Relationships between nodes in social and acoustic similarity networks

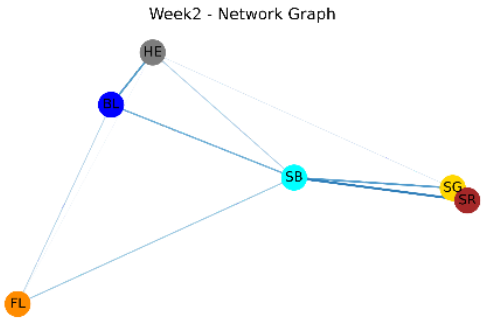
Finally, we tested whether individual dominance rank is related to overall acoustic similarity in the network. Our previous tests measured similarity between dyads as a function of rank difference, but we here measured whether rank dictated distance from the centre of the overall acoustic network. As shown in Figure 18, we measured correlations between each bird's dominance rank (David's score) with its normalized weighted degree in an "acoustic similarity" network, where edges represent acoustic similarity based on distance or counts.

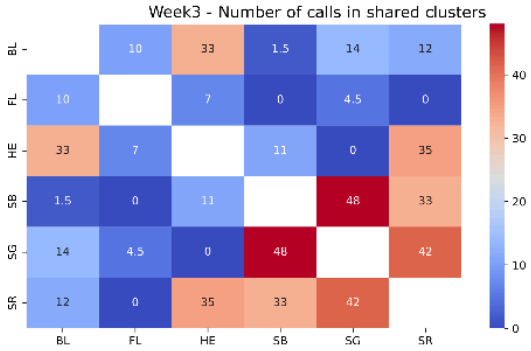
Across all our metrics, all measures of call subtype usage (and only this category of metric) were significantly correlated with rank as shown in Figure 19. We found a negative correlation between dominance rank and the number of calls shared in a cluster (rho = −0.57, p = 0.003, z = −2.682) as well as the number of shared clusters (rho = −0.44, p = 0.01, z = −2.133). Therefore, higher-ranked birds appeared to occupy more central positions in the call-subtype similarity network, and birds closer to them incorporated more similar call subtypes in their warbles.

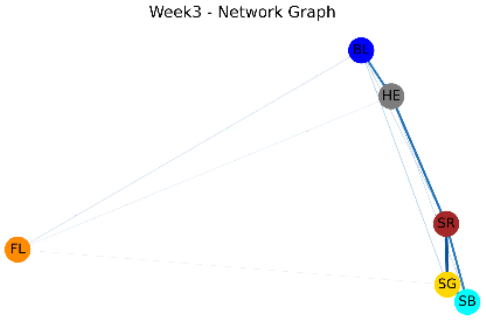

Week1 - Number of calls in shared clusters
BL
FL
HE
SB
SG
SR
4 7.5 17 0 0
4 16 7 10 6
7.5 16 22 0 5.5
17 7 22 5.5 21
0 10 0 5.5 49
0 6 5.5 21 49
0 10 20 30 40

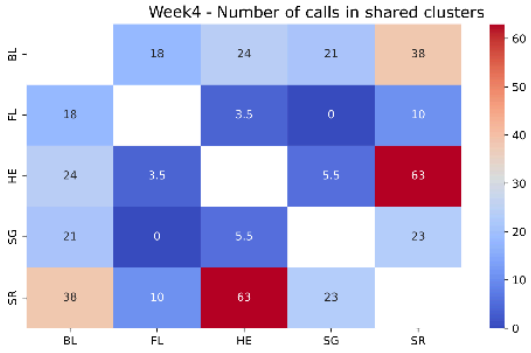

Week1 - Network Graph
SG
SR
BL
SB
HE
FL

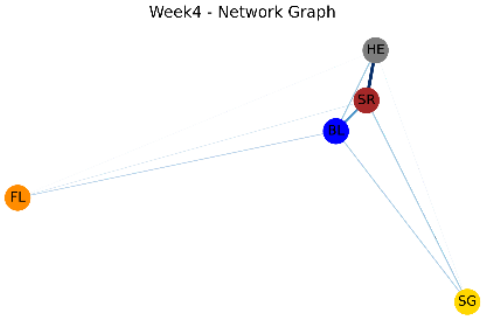

Week2 - Number of calls in shared clusters
BL
FL
HE
SB
SG
SR
18 36 30 0 0
18 5.5 24 0 0
36 5.5 20 10 0
30 24 20 36 44
0 0 10 36 62
0 0 0 44 62
0 10 20 30 40 50 60


Week2 - Network Graph
HE
BL
SB
SG
SR
FL


Week3 - Number of calls in shared clusters
BL
FL
HE
SB
SG
SR
10 33 1.5 14 12
10 7 0 4.5 0
33 7 11 0 35
1.5 0 11 48 33
14 4.5 0 48 42
12 0 35 33 42
0 10 20 30 40

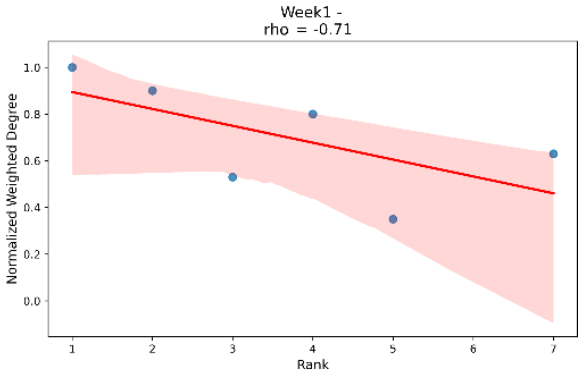

Week3 - Network Graph
BL
HE
SR
FL
SG
SB


Week4 - Number of calls in shared clusters
BL
FL
HE
SG
SR
18 24 21 38
18 3.5 0 10
24 3.5 5.5 63
21 0 5.5 23
38 10 63 23
0 10 20 30 40 50 60


Week4 - Network Graph
HE
SR
BL
FL
SG

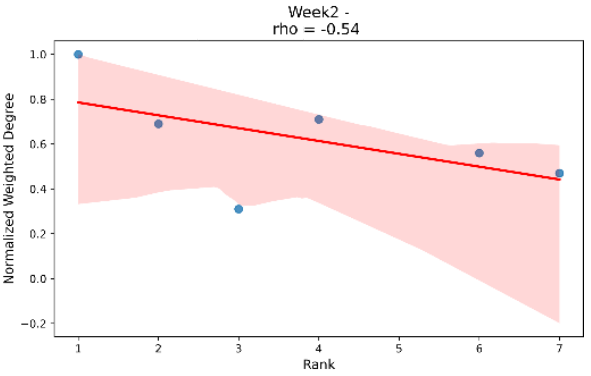
Week1 - Metrics

Bird	NWD	David's Rank
BL	0.35	5.0
FL	0.53	3.0
HE	0.63	7.0
SB	0.9	2.0
SG	0.8	4.0
SR	1.0	1.0



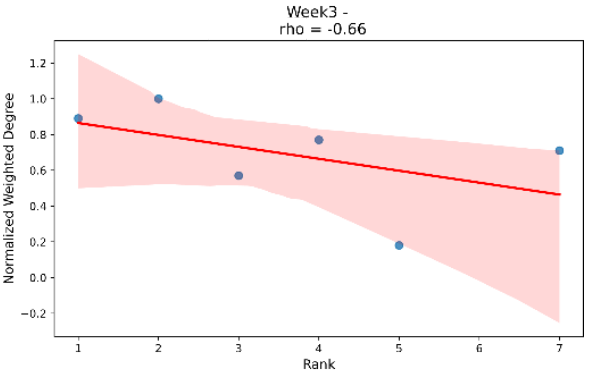
Week2 - Metrics

Bird	NWD	David's Rank
BL	0.56	6.0
FL	0.31	3.0
HE	0.47	7.0
SB	1.0	1.0
SG	0.71	4.0
SR	0.69	2.0



Week3 - Metrics

Bird	NWD	David's Rank
BL	0.57	3.0
FL	0.18	5.0
HE	0.71	7.0
SB	0.77	4.0
SG	0.89	1.0
SR	1.0	2.0



Week4 - Metrics

Bird	NWD	David's Rank
BL	0.75	5.0
FL	0.24	7.0
HE	0.71	6.0
SG	0.37	2.0
SR	1.0	1.0

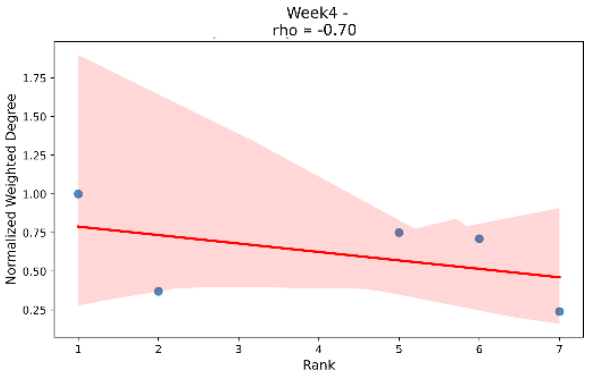


Figure 17 *Example of an acoustic similarity network constructed using the call-distance matrix. From this, we derived a node's normalized weighted degree as a measure of how similar that individual's calls were to all others. The node-level measure was then compared to rank in the social network-derived dominance hierarchy (from David's score).*

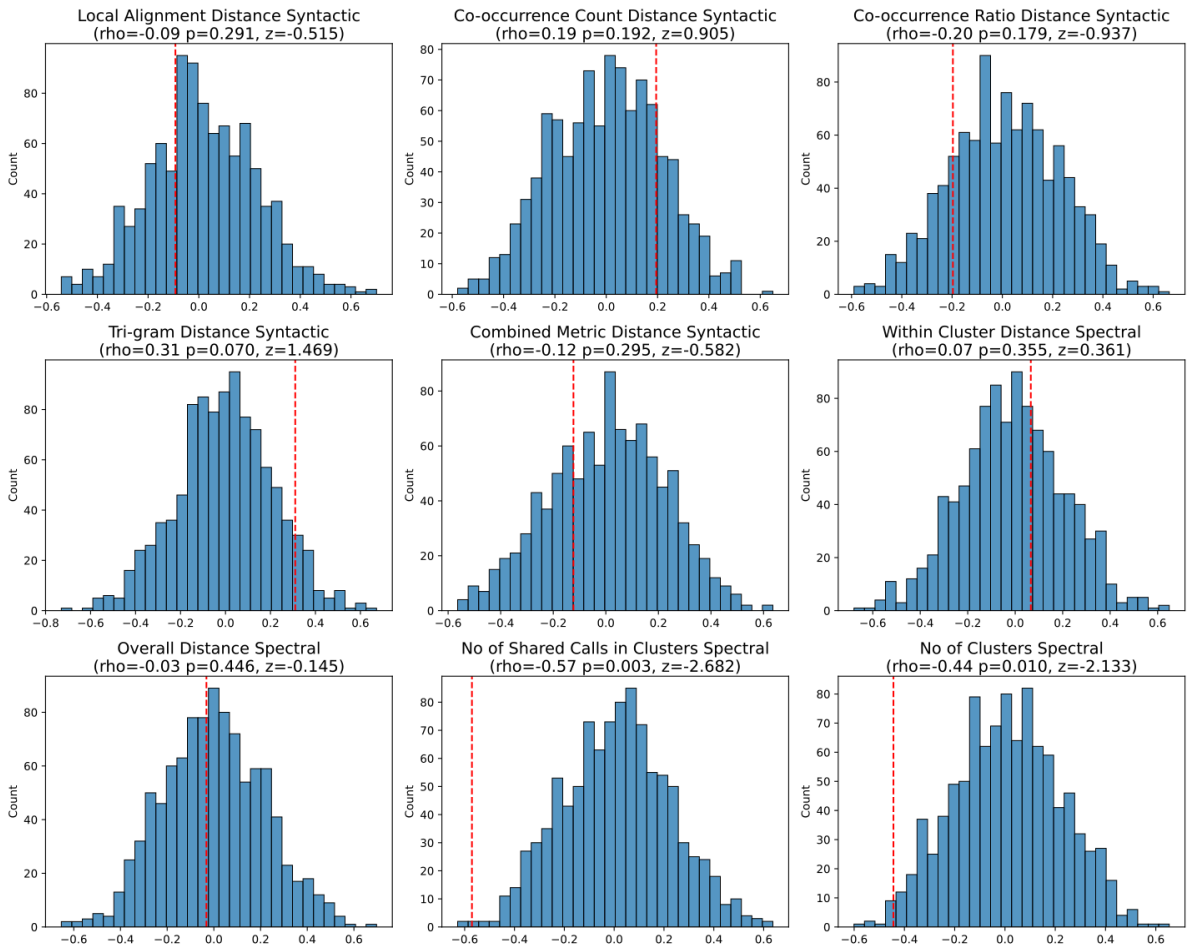


***Figure 18**: Distribution of correlation coefficients from permuted null datasets, where each bird's normalized weighted degree in the acoustic similarity network is compared to its dominance rank. The observed correlation is marked by a vertical red dotted line. The permutation-based p-value and z-score are shown in the plot title.*

Chapter 4: Discussion

4.1 Social structure

We observed that budgerigars form a clear social structure, where aggressive interactions are organized into a dominance hierarchy. Females consistently outrank males, and although the male hierarchy is somewhat steeper than that of females, both remain relatively shallow overall and exhibit relatively little change over the course of four weeks. Interestingly, the same display behavior that males direct toward females is also used in male–male interactions, where it appears to function as a submissive signal. The occurrence of male-male displays is negatively correlated with directed aggression. These findings extend the limited body of previous knowledge about parrot social systems, providing a detailed analysis of dominance relationships and the role of male-male displays in these hierarchical relationships.

4.1.1 Why do parrots have dominance hierarchies?

Dominance hierarchies establish a rank order, thereby minimizing the need for direct aggression. Although aggression is more common when ranks are being settled, the final hierarchy is largely maintained through threats, submissive signals, and occasional punishments by dominant individuals if subordinates challenge their status (Tibbetts et al., 2022). However, repeated interactions within a consistent set of group members are essential for a stable hierarchy to develop. In parrots, where fission-fusion dynamics often cause large flocks to split and re-form, recent findings indicate that stable subgroups persist over long periods, permitting the emergence of stable rank relationships (Bradbury & Balsby, 2016; Wright & Dahlin, 2018). For example, studies on sulphur-crested cockatoos *(Cacatua galerita)* demonstrated long term social associations between individuals and stable dominance structures within these subgroups (Penndorf et al., 2024; Penndorf, Ewart, et al., 2023; Penndorf, Farine, et al., 2023). Parallel evidence from other parrot species also points to long-term affiliative associations, potentially guided by individual recognition and contact-call discrimination (Berg et al., 2011; Wanker et al., 1998, 2005). For

example, Orange-fronted conures *(Aratinga canicularis)* approach and respond more vigorously to calls from locally familiar birds, suggesting that they can distinguish between in-group and out-group members (Vehrencamp et al., 2003). Consequently, when subgroups remain consistent enough to foster repeated encounters, the formation of a dominance hierarchy is expected, even amidst the broader fission-fusion social dynamics characteristic of parrots.

4.1.2 Females are higher ranked than males in the dominance hierarchy

We observed that females consistently outranked males, even when we excluded incidents in which females displayed aggression following a courtship attempt. This observation agrees with earlier qualitative reports describing higher female aggression in budgerigars (Abbassi & Burley, 2012; Brockway, 1964a, 1964b). However, female dominance is unusual among parrots; for instance, males dominate in sulphur-crested cockatoos *(Cacatua galerita)* and cockatiels *(Nymphicus hollandicus)*, whereas sex does not reliably predict rank position in monk parakeets (Hobson et al., 2014; Penndorf et al., 2024; Seibert & Crowell-Davis, 2001). Males typically hold dominant status in many other bird taxa (e.g., crows *(Corvus macrorhynchos)*, chickens *(Gallus gallus domesticus)*, and various songbirds) (Kondo & Hiraiwa-Hasegawa, 2015). Nonetheless, although female dominance has not previously been reported for parrots, it is known in mammals such as brown hyaenas *(Hyaena brunnea)* and African elephants *(Loxodonta africana)* (Owens & Owens, 1996; Wittemyer & Getz, 2007).

Budgerigars generally retained their positions within the dominance hierarchy over the course of four weeks, with only minor shifts involving single-rank changes. This pattern parallels findings in monk parakeets, where established, hierarchical ranks remain stable over time (Hobson & DeDeo, 2015; Van Der Marel et al., 2023). Taken together, our data suggest that budgerigar dominance relationships, once formed, exhibit temporal consistency at least under captive conditions.

4.1.3 Shallow dominance hierarchies in parrots

We found significant negative reciprocity in aggression between dyad members: the recipient of aggression rarely reciprocated, consistent with a subordinate–dominant dynamic. Nevertheless, the hierarchies we observed in both sexes were relatively shallow and less linear than those in crows *(Corvus macrorhynchos)* or chimpanzees *(Pan troglodytes)*, where hierarchy nearly always predicts the outcome of aggressive interactions (Kondo & Hiraiwa-Hasegawa, 2015; Tibbetts et al., 2022; Watts, 2018). By contrast, the lower steepness and linearity in budgerigars indicate that aggression outcomes are not fully determined by rank alone; other factors may thus also influence the success or failure of a challenge. Our findings align with those reported for other parrot species (monk parakeets *(Myiopsitta monachus)*, sulphur-crested cockatoos *(Cacatua galerita)*, and cockatiels (*Nymphicus hollandicus*)) (Hobson et al., 2014; Penndorf et al., 2024; Seibert & Crowell-Davis, 2001). Specifically, budgerigars exhibit a steeper hierarchy than monk parakeets but a shallower one than sulphur-crested cockatoos. It is important to emphasize that low steepness does not negate the existence of a clear rank order. For example, monk parakeets display even lower steepness, yet show hierarchical aggression patterns consistent with a transitive, dominance-based system. Here, aggression is mostly directed at individuals of adjacent rank, suggesting that subordinates use transitivity to evaluate their own status relative to others (Hobson & DeDeo, 2015).

The relatively modest steepness values observed in budgerigars and other parrots are unsurprising when one considers their social structure and flocking patterns. In group-living species with a stable composition such as crows *(Corvus macrorhynchos)* or chimpanzees *(Pan troglodytes)*, subordinates rarely challenge dominants. Although lower-ranked individuals gain fewer resources and mating opportunities in such systems, they benefit from stable group membership (e.g., foraging efficiency, and predator vigilance). Parrots, however, exhibit fission-fusion flocking, and a subordinate bird could, in theory, leave one group and join another where it might achieve a more favorable rank. The turnover introduced by such movements and the subsequent need for newcomers to negotiate or reestablish rank could foster more frequent challenges. Indeed, in newly formed monk parakeet *(Myiopsitta monachus)* groups, steepness increases, reaching a plateau over time as

hierarchies stabilize (Hobson & DeDeo, 2015). In natural contexts, ongoing group reshuffling may persistently undermine steepness by enabling subordinates to either challenge higher-ranked individuals or depart if upward mobility seems limited. Although direct evidence linking fission-fusion dynamics to shallower hierarchies remains elusive, our results support the hypothesis that parrots, by virtue of their fluid social structure, form relatively shallow hierarchies compared to taxa with more stable group membership. However, as our studies were carried out on a captive colony where this process does not naturally occur, more studies in the wild or experimental translocations within captive colonies are needed to test these hypotheses, which must at this point remain speculative.

4.1.4 Male-male displays as signals of submission

Our analysis revealed a negative correlation between the number of courtship-like displays and aggressive interactions within male-male dyads. Specifically, when one male displayed more to another, the aggression it received from that male was significantly lower than expected by chance. Moreover, a recipient of aggression was more likely to display to its aggressor soon after than expected by chance. These observations suggest that in male-male contexts, displays serve as an appeasement or submissive signal, helping to mitigate conflict without resorting to overt aggression. Notably, male-male displays were at least as common as male-female displays, indicating that these interactions may be crucial to maintaining the dominance hierarchy.

An alternative hypothesis is that such displays function as “practice” to improve courtship skills, particularly in younger males (Abbassi and Burley, 2012). According to this idea, males rehearse their displays on other males to reduce the risk of aggression from females, which may be more aggressive toward poor-quality performances. However, contrary to the prediction that practice would improve future mating success, Abbassi and Burley (2012) found that males who performed fewer male–male displays secured mates more readily. They therefore suggested that male–male displays might instead be a strategy for subordinate males to demonstrate vigor or stamina. More physically fit individuals who have already established their physical vigor in direct aggressive contests display less frequently

and, in turn, tend to achieve greater success with females. Although this study did not directly measure aggression and thus could not infer dominance relations among the males, the combined evidence aligns with our hypothesis of displays as a submissive gesture. Higher-ranking males, with less need to signal submission, display less often, but are better at securing mates. The resultant improved access to mates is consistent with other benefits of high social rank, such as prioritized access to resources. Dominant individuals also feed before subordinates, highlighting the dual importance of rank in both resource acquisition and mate choice (Abbassi & Burley, 2012; Hasegawa & Soma, 2004).

In primates, steep (or “despotic”) hierarchies rely primarily on dominant and aggressive signals to uphold rank, whereas shallower (or “egalitarian”) hierarchies (where outcomes are not entirely determined by rank) often feature more frequent displays of submission to reduce conflict. Budgerigars, which exhibit relatively shallow dominance hierarchies, thus align with the pattern expected under an “egalitarian” hierarchy and employ submissive signals to maintain social stability (Girard-Buttoz et al., 2022; Maciej et al., 2013; Preuschoft et al., 2001).

Our criterion for defining a “display” remained consistent with previous usage, namely any sequence of display-like body movements (Wewhare & Krishnan, 2024). It is possible, however, that the composition and sequencing of male-male displays differ from those directed toward females, an intriguing avenue for future research. For example, in male Jacky dragons *(Amphibolurus muricatus)*, the different movement components convey contrasting signals, as “push-up displays” are an aggressive signal and “slow arm-waves” are a submissive signal (Van Dyk & Evans, 2008).

Although captive studies yield valuable mechanistic insights, complementary observations in the wild are essential to confirm both the existence and role of dominance hierarchies under natural conditions. The only available field study (Wyndham, 1980) employed qualitative methods, observing low aggression in wild budgerigars but higher aggression in cage birds. From these results, the author speculated that dominance hierarchies might be an artifact of captivity, where ample food, confined spaces, and limited activities increase aggression. However, this

conclusion remains tentative because the observer did not track individual birds or systematically quantify aggression. A more robust approach would involve systematically varying cage size and resource distribution while controlling for group size, allowing researchers to assess how these factors influence dominance relationships. If, as Wyndham posited, abundant food leads to increased aggression, one might see a stronger dominance signature when resources are plentiful. In contrast, classical socioecological theory suggests that resource scarcity enhances competition, thereby reinforcing hierarchies (Tibbetts et al., 2022). Testing these opposing predictions would clarify whether dominance hierarchies in budgerigars arise primarily from captive artifacts or reflect an intrinsic feature of their social organization.

4.2. Individual-specific acoustic signatures

4.2.1 Individual-level signatures in warble syntax

We uncovered clear evidence that budgerigars which had lived together in the same colony for at least three months maintained individual signatures in their warble syntax. However, there was also a statistically significant temporal drift over just a few weeks. Thus, although each individual maintains a distinct style of structuring its warble compared to others, this unique style also changes over time. Unlike many songbirds whose highly consistent syllable ordering can often be modeled by finite-state grammars or first-order Markov chains, budgerigar warble sequences are more variable both in length and content, even within an individual. Consequently, our analyses were not performed under the expectation that sequences would be wholly identical or entirely distinct (Kershenbaum et al., 2014, 2016; Madabhushi et al., 2023). Rather, we detected segments of similarity within longer sequences. Because no two warbles are ever exact replicates even within the same individual, our metrics (e.g., local alignment, tri-gram analysis, and co-occurrence) focused on identifying these shared regions across otherwise heterogeneous portions of song.

Earlier work by Madabhushi et al., (2023) demonstrated that when previously separated budgerigar flocks merge, their syntax (for example, the length of repeated

motifs of certain notes) converges, thus encoding group-level identity. Our findings extend this idea by showing that within a single colony, each bird also maintains a distinctive syntactic signature. At the same time, individuals are more similar to each other than to birds from entirely different groups. Historical field observations on budgerigars likewise suggest that warble song may facilitate individual recognition and serve additional functions, such as attracting females to nest holes when males return from foraging (Brockway, 1964a; Wyndham, 1980).

Our study provides the first demonstration of individual-level syntactic signatures in parrot vocalizations. In meerkats (*Suricata suricatta*), sentinel calls possess call-type orders that remain consistent within individuals, whereas rock hyraxes (*Procavia capensis*) use multiple acoustic features in combination, resulting in male songs that differ in element sequencing between individuals (Rauber et al., 2020). In domesticated canaries *(Serinus canaria domestica)*, each bird's repertoire mixes widely shared syllables with unique sequences of three or more syllable types, creating cues of individual identity (Lehongre et al., 2008). Neighboring male skylarks *(Alauda arvensis)* share particular phrase sequences, whereas distant strangers have none in common, enabling territory owners to discriminate local birds from new intruders (Briefer et al., 2008; Briefer et al., 2013). Additional examples come from chickadees *(Poecile atricapillus)*, which rely on note-order differences to recognize individuals in their flocks, and from cetaceans, where patterned sequences e.g., "codas" in sperm whales *(Physeter macrocephalus)* may serve as signatures of identity (Krams et al., 2012; Oliveira et al., 2016). The songs of budgerigars are longer, more variable, and more complex than most of the species discussed above, highlighting the fact that individual identity cues operate across a broad range of songs with varying complexities. Taken together, these findings situate our budgerigar results within a broad cross-species context, showing that individual-specific vocal syntax is a widespread and integral component of animal communication.

4.2.2 Individual signatures in spectral properties of contact calls and call subtype use

Unlike syntax, budgerigar contact calls (or call-like elements in warble song) have received more study (Bartlett & Slater, 1999; Dahlin et al., 2014; Farabaugh et al., 1994; Hile & Striedter, 2000). In line with this previous research, our spectral analyses revealed that these calls are grouped into multiple subtypes (Tu et al., 2011; Zhao et al., 2023). Each individual exhibits a distinct signature both in the spectral properties of each subtype emitted by it, and in the relative proportions of different subtypes in its repertoire. Consequently, individuals remain distinguishable from one another, even within the same subtype cluster. Similar to patterns in syntax, we also observed temporal drift in the overall call repertoire on a weekly timescale, driven primarily by changes in the relative usage of different subtypes rather than shifts in the spectral properties of any one subtype. In other words, the spectral properties of different contact calls remain stable across weeks, whereas the proportion of each subtype in the repertoire changes.

4.2.3 Multimodal signaling of individual identity

Identity signals are generally predicted to be highly variable across individuals, yet relatively stable within individuals, not linked to condition or fitness, and often composed of multiple independent cues (Dale et al., 2001). Our findings in budgerigars across multiple studies align with these principles: each bird displays distinct spectral patterns in contact calls, individual-specific syntactic structures in warble sequences, and unique display-based gestures (Dahlin et al., 2014; Madabhushi et al., 2023; Wewhare & Krishnan, 2024). Together, these channels form an integrated system to reliably convey identity, echoing the broader observation that misidentification may impose considerable costs in dense or socially complex groups (Dale et al., 2001). By employing multiple cues (none of which appear tightly coupled to the condition) budgerigars illustrate how identity signals may evolve when social recognition plays a central role in mediating various interactions. Moreover, our results are consistent with the idea that, in species requiring robust individual recognition, evolutionary pressures favor multimodal or hierarchically organized traits that preserve individuality despite group-level

convergence or drift in any single dimension (Dale et al., 2001; Tibbetts & Dale, 2007).

This idea has found broad support across other taxa, including colonial swallow chicks, whose begging calls vary dramatically among offspring to ensure accurate parental care, or ruffs and queleas that display conspicuously variable plumage to signal individual identity (Medvin et al., 1993; Tibbetts & Dale, 2007). In humans, facial diversity may serve a similar function: large group sizes and frequent social interactions make it beneficial for individuals to be highly recognizable via multiple traits (Jack & Schyns, 2015).

4.3 Vocal signal similarity is influenced by the social network

4.3.1 Social structure and female choice influence call subtype use

We found that males which (in the colony) directed more displays toward the same females, and exhibited more call subtype sharing in their warble songs than expected by chance. By contrast, males that displayed more to each other differed significantly in their call subtype usage. Prior literature indicated that male budgerigars learn and repeatedly produce the contact calls of the females they court. Thus, when two males display to the same female(s), they both tend to adopt call features that approximate her call structure, thus converging on the same subtype (Hile et al., 2000; Moravec et al., 2006). Although these studies were conducted in simple one-male–one-female paradigms, our findings extend these findings to a multi-male, multi-female social network. Earlier studies also found that female budgerigars prefer males whose contact calls resemble their own, and our findings broadly support this idea. Divergence in call type use among males that display more to each other could be because they court different females and thus diverge in contact call subtypes as they attempt to become similar to different females. However, this is speculative and has not been tested.

Although males with similar patterns of interactions with females converge on similar contact call subtypes, they simultaneously diverge (greater than expected by chance) in warble-song syntax. In other words, the very same dyads that align in call subtype usage differ in warble structure. One potential explanation is that each male needs to be distinguished from another displaying to the same female in some aspect of vocal performance namely, by varying note ordering, repetition patterns, or phrase sequences. For example, in canaries *(Serinus canaria domestica)*, although males share parts of their syllable repertoire with others, they arrange them in unique combinations, such that most sequences of more than three syllables are individually unique. These unique, individual-specific combinations occur frequently in the song of a given individual and thus might enable vocal recognition without requiring knowledge of its full repertoire (Lehongre et al., 2008). And in monk parakeets *(Myiopsitta monachus)* individuals with stronger affiliative bonds have more dissimilar calls, indicating an active attempt by individuals to sound unique among close associates (Smeele et al., 2024). Alternatively, the female's variable responses could shape each male's warble structure. A clear preference for one male may lead to differences in display behaviors between the two males. If the female favors one male over the other, she may allow him to allopreen and allofeed and deny these interactions to the other male. As a result, the preferred male could incorporate allopreening and allofeeding into his body movement sequence during display, whereas the other male cannot. This leads to distinct body movement sequences when both males court the same female. Because each body movement is closely tied to specific notes in the warble, these behavioral differences during display could ultimately drive divergence in their song syntax (Wewhare & Krishnan, 2024).

We note here, however, that although the colony's social behaviors were recorded in a shared enclosure providing robust data on who displayed to whom, the high noise level within the colony prevented us from capturing individual vocalizations in that same setting. Consequently, we recorded vocal data in a separate chamber, pairing each male with the female he displayed most often. This design introduces the possibility that the observed vocal patterns may have been influenced by individual variation in each female's responses during these controlled recordings. However, our analytical approach, focusing on whether similarity or divergence exceeds

chance levels should mitigate this concern, because it compares inter-individual vocal patterns to a robust null model. Further, the possible presence of individual variation in female responses actually strengthens the point we make above, that of female choice driving convergence or divergence in the song of male dyads. In any case, future studies incorporating individual-level vocal recordings from within the colony itself would offer clearer insights into how female feedback directly shapes vocal repertoires in more naturalistic contexts.

A similar pattern of context-driven vocal convergence has been documented in other taxa, particularly primates. For instance, Marmosets (*Callithrix jacchus*), convergent "trill" calls indicate social closeness, but calls crucial for individual recognition (e.g., "phee" calls) remain more stable (Zürcher et al., 2021). In Female Diana monkeys (*Cercopithecus diana*) actively match the frequency contours of preceding calls during direct vocal interactions, presumably to maintain spatial proximity and social cohesion (Candiotti et al., 2012). Among birds, the basal passerine rifleman (*Acanthisitta chloris*) and chickadees (*Poecile atricapillus*) both exhibit rapid call convergence when individuals form close associations (Krams et al., 2012; Moran et al., 2024). Although the specific social drivers range from pair bonding to group stability, the consistent thread across these diverse systems is that targeted vocal accommodation can strengthen affiliative bonds.

4.3.2 Position in the dominance hierarchy influences individual vocal repertoires

Finally, our study found that budgerigars located at higher rank in the hierarchy are more centrally located in an acoustic network based on contact call subtype usage. This implies that all birds tend to align their call types with those of more dominant flock members in a rank-dependent manner. Moreover, the central position of dominant individuals in the contact call network suggests that each subordinate bird converges on the call usage of those above it in rank, potentially as an affiliative gesture to signal submission and thus prevent aggression. This, although speculative at this point, presents an important direction for future study.

We also found that males positioned closer in rank within the dominance hierarchy exhibit greater similarity in their warble-song syntax than expected by chance. One possible explanation for this pattern is that near-ranked individuals frequently display to each other while negotiating dominance, leading them to learn and adopt each other's note arrangements. Alternatively, these males may also learn each other's patterns of body movements during displays, which could, in turn, shape parallel syntactic structures, as warble notes are closely tied to physical gestures (Wewhare & Krishnan, 2024).

Our study is the first to directly link vocal convergence with the mediation of dominance hierarchy, despite extensive separate research on both rank-driven social structure and call convergence in parrots (Bradbury & Balsby, 2016; Hobson, 2022). A related phenomenon occurs at the group level in orange-fronted conures (*Aratinga canicularis*): when two flocks prepare to merge, a series of contact call exchanges occurs before the merger where one group converges on the other's contact calls, effectively becoming the "follower." Playback experiments further showed that bystander flocks eavesdrop and identify which flock is likely to become the leader, indicating that convergence is a robust affiliative cue (Thomsen et al., 2021). In other taxa, vocal convergence is also hypothesized to serve as an affiliative signal. Both sperm whales (*Physeter macrocephalus*) and Japanese macaques (*Macaca fuscata*) actively adjust their calls to match the vocalizations they hear from other individuals, ostensibly to signal affiliation. Sperm whales exchange short click sequences called "codas," often responding with the same coda type in a duet-like chain, thereby reinforcing social bonds even among whales separated by a considerable distance (Schulz et al., 2008). Similarly, Japanese macaques tested in playback experiments matched the acoustic characteristics of the calls they heard, indicating that vocal convergence in "coo" calls may help maintain contact and cohesion (Sugiura, 2007). Thus, although there is evidence of call as an affiliative signal, our study is the first to match it to a dominance-based social network, and we propose that call convergence also serves to mediate and reinforce hierarchies within groups. With the evidence above, studies of other taxa using consensus among a broad range of metrics may uncover a similar pattern.

4.3.3 Open-ended vocal learning in a complex social setting

Our findings underscore how multifaceted social interactions shape the vocal repertoire of budgerigars independently at multiple hierarchical levels. On one hand, males converge in contact call usage when courting the same female, yet they simultaneously diverge in their warble syntax. On the other hand, individuals appear to match the vocal patterns of a dominant bird in a rank-dependent manner. These seemingly opposing patterns: convergence in one vocal domain and divergence in another, illustrate the remarkable social and vocal flexibility of budgerigars. Social structure itself can shift rapidly: both dominance hierarchies and courtship affiliations may reorganize over time, and in fission-fusion contexts similar to those in the wild, individuals regularly move among different groups, adding further layers of complexity.

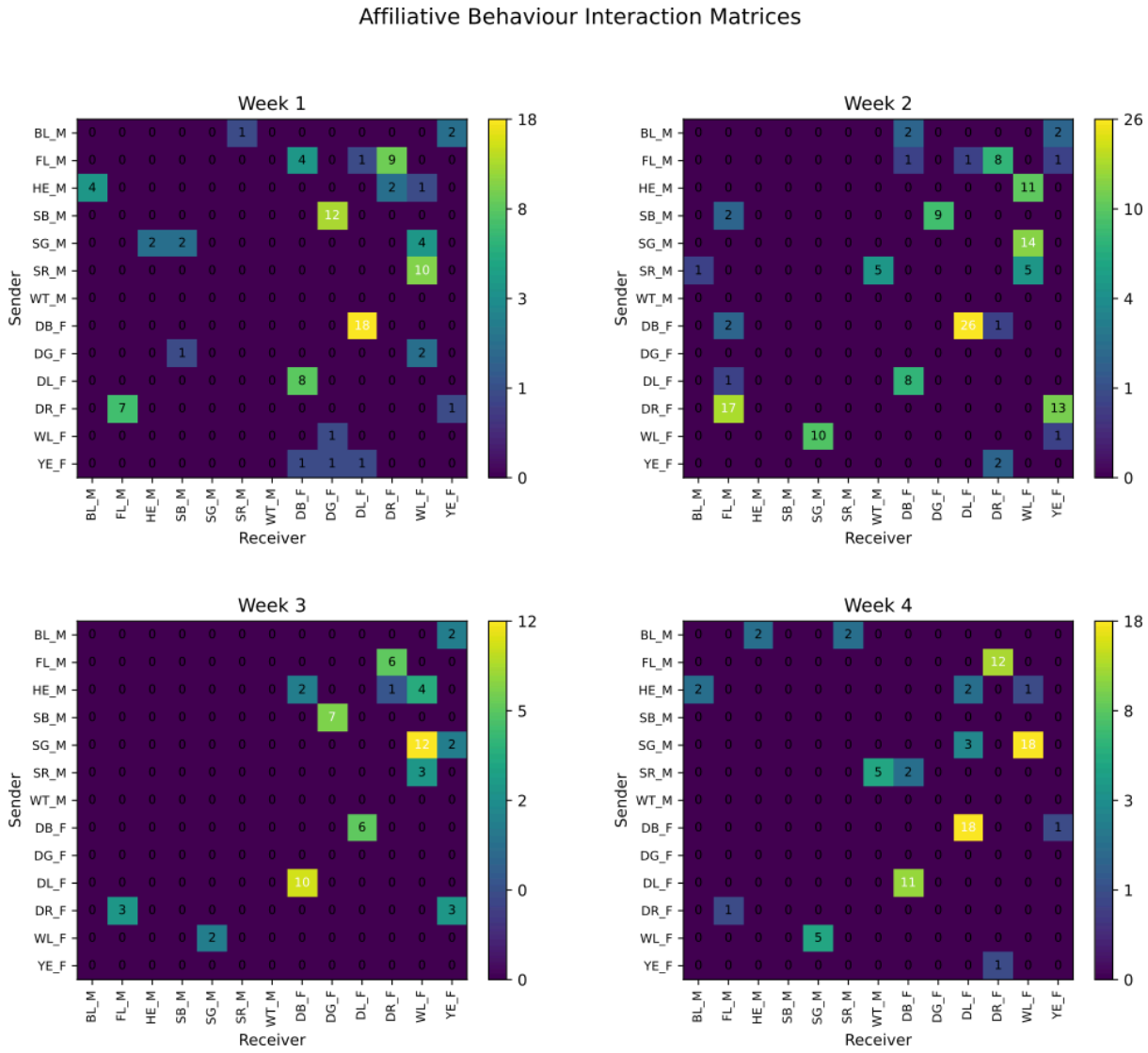
Because budgerigars occupy a highly fluid social environment, their vocal systems must be correspondingly dynamic, capable of adjusting across multiple acoustic axes: contact call structure, contact call subtype usage, warble syntax, and perhaps other unstudied axes of variation to accommodate changing dyadic interactions, hierarchies, and mate preferences. Our findings broadly align with the “social complexity hypothesis,” which posits that as a species’ social structures become more intricate, so too do their communicative repertoires (Bouchet et al., 2013; Krams et al., 2012; Smeele et al., 2024). Parrots, with their tendency for dynamic, fission-fusion group organization even at very short timescales, require signals that can both identify a shared group identity and handle stable subgroup dynamics, including dominance. Thus, our results reinforce the idea that budgerigar's open-ended vocal learning manifests in highly adaptable call usage and syntax, which plays a critical role in maintaining cohesion and managing social relationships in an ever-shifting social landscape.

Chapter 5: Conclusion

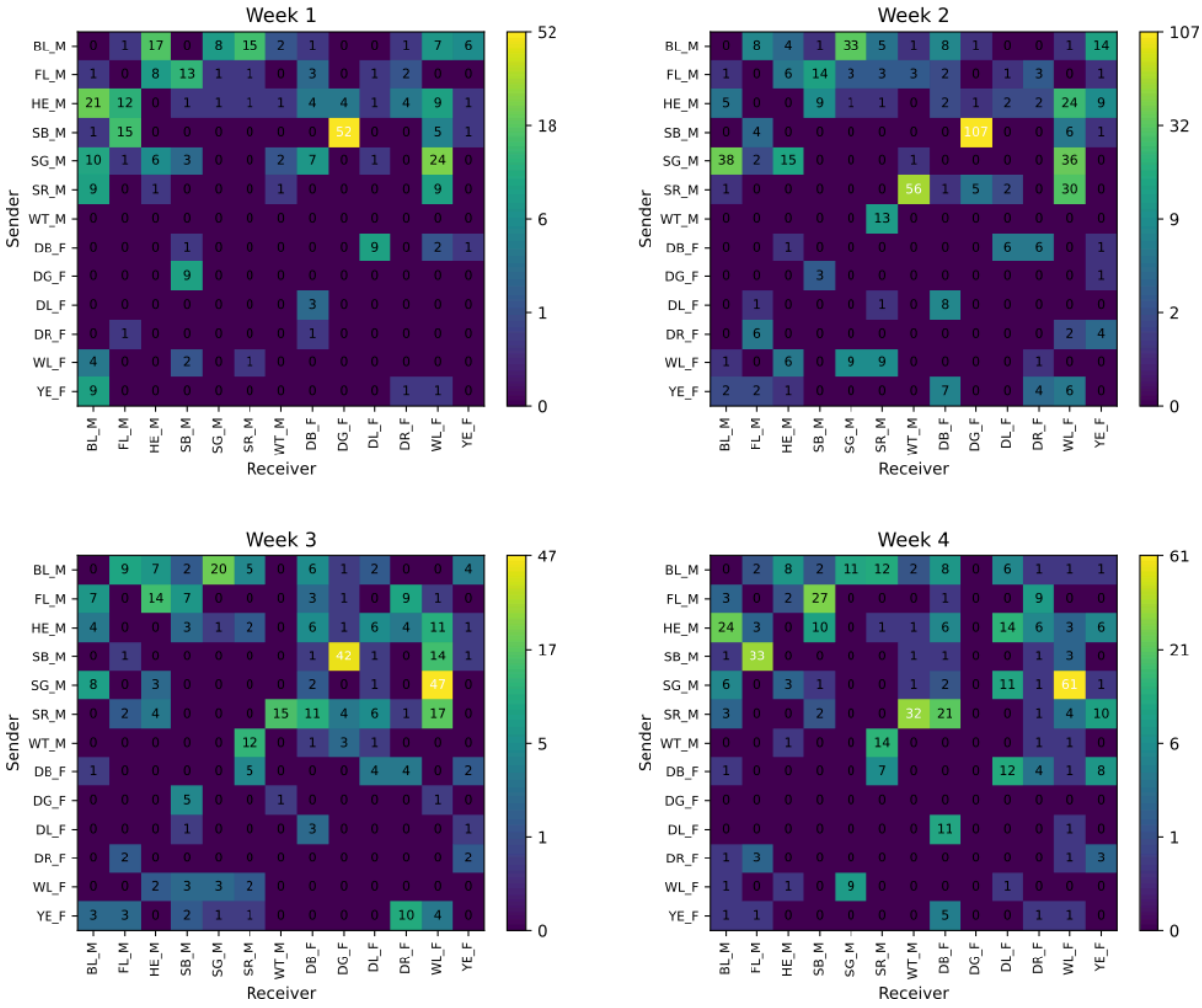
Our investigation of budgerigar social structure and vocal behavior reveals how multiple, seemingly contrasting pressures shape group living. First, we show that budgerigars form stable if relatively shallow, dominance hierarchies in which females consistently outrank males. Once established, these hierarchies remain largely unchanged over several weeks. Observing that a display commonly used in courtship also appears in male-male contexts highlights the dual function of display behavior: beyond courtship; it can serve as a submissive signal that helps subordinates avoid overt aggression.

Second, our acoustic analyses underscore how budgerigars integrate individual recognition, mate choice, and rank negotiation within a highly flexible vocal communication system. Each bird's warble syntax bears an individual signature and yet can drift in just a few weeks. Contact calls are similarly labile: individual spectral features stay consistent enough to mark identity, but males may converge in contact call subtypes and diverge in syntax if they court the same female. Crucially, we also find that birds higher in the dominance hierarchy occupy a "central" position in the call-subtype network, with subordinates exhibiting convergence towards these individuals in a rank-dependent manner. This acoustic centrality suggests the possibility that subordinate birds actively signal rank-dependent affiliation via vocal matching. Stable subgroups or repeated interactions enable hierarchies to persist even under fission-fusion conditions typical of wild parrots. At the same time, open-ended vocal learning enables quick shifts in contact calls or warble syntax to navigate changing alliances, courting opportunities, and rank negotiations. The amenability to laboratory manipulations thus renders the budgerigar a powerful study system in socioecology and social learning.

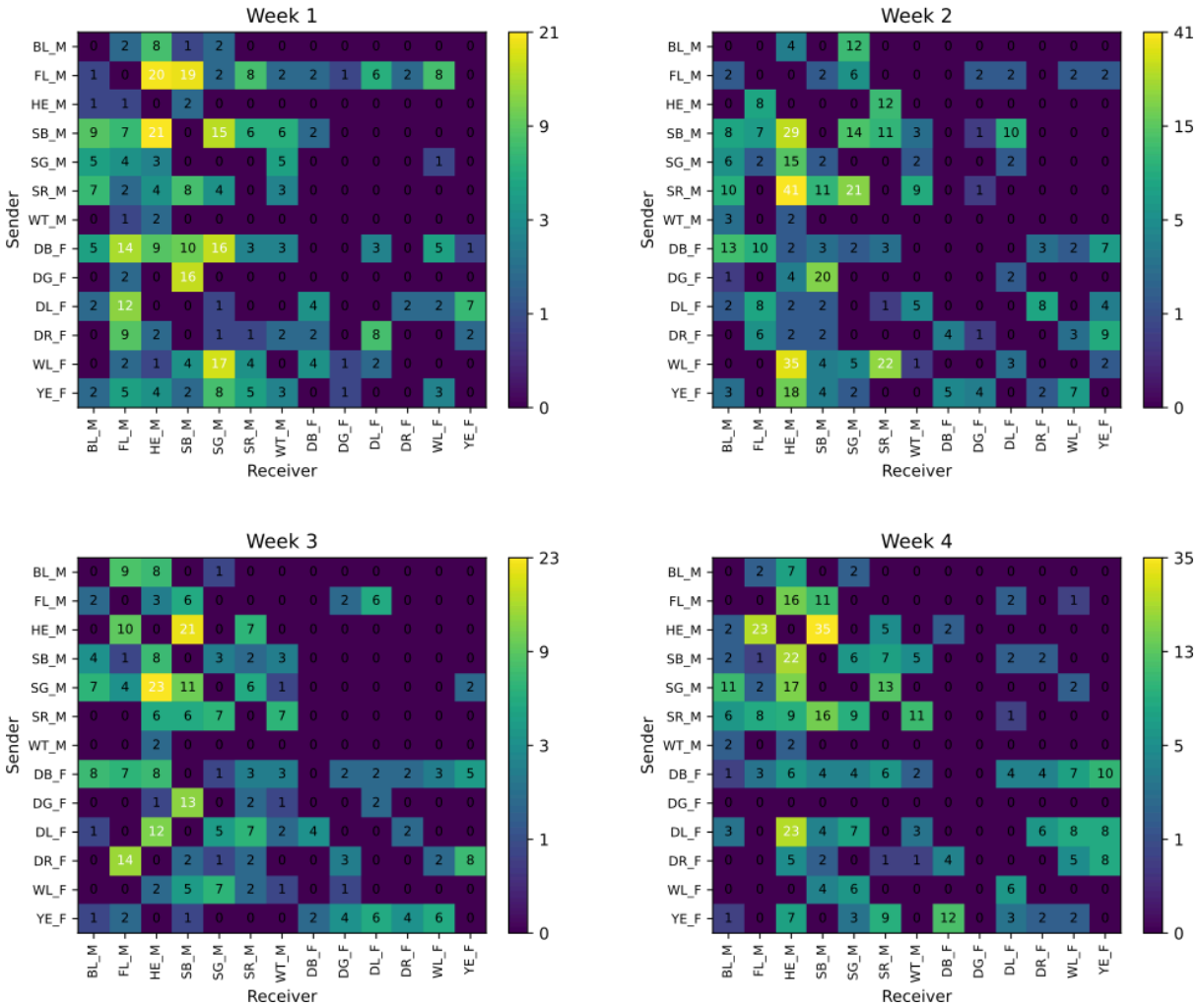
Appendix

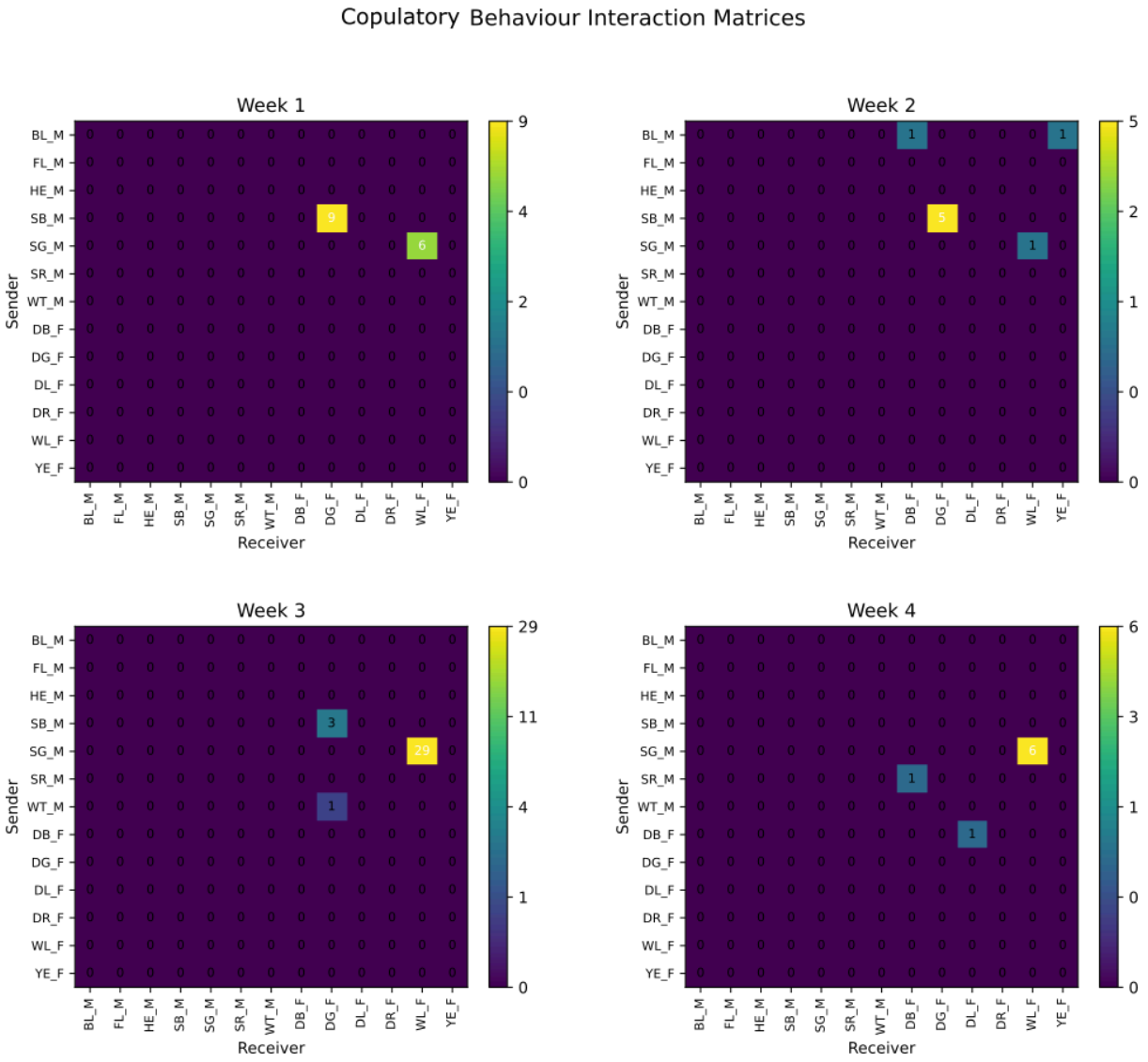

Affiliative Behaviour Interaction Matrices
Week 1
Week 2
Week 3
Week 4
Sender
Receiver
BL_M
FL_M
HE_M
SB_M
SG_M
SR_M
WT_M
DB_F
DG_F
DL_F
DR_F
WL_F
YE_F

Display Behaviour Interaction Matrices



Aggression Behaviour Interaction Matrices





***Supplementary figure 1**: Behavioral interaction matrix per behavior category per week, visualized as heatmaps. This represents the raw data that will be used for all further analyses. (A) Affiliative, (B) Display, (C) Aggression, (D) Copulatory.*

Behavior Metric	Acoustic Metric	Observed Slope	Slope Z-score	Slope P-value	Observed Spearman Rho	Rho Z-score	Rho P-value
Display	local alignment	-0.00034	2.80472	0.003	-0.11542	0.885525	0.182
Display	co-occurrence count	-0.00231	-1.97613	0.02	0.038499	0.100875	0.456
Display	co-occurrence ratio	-8.04E-05	-2.22139	0.01	0.064613	-1.87574	0.033
Display	tri-gram	-0.00057	-1.69041	0.043	0.171101	2.318902	0.009
Display	combined metric syntax	-0.00091	-1.57259	0.058	-0.00883	-1.97064	0.023
Display	within cluster distnace	8.17E-06	-1.5844	0.047	0.447965	4.236723	<0.001
Display	overal spectral distance	7.26E-06	-3.08035	<0.001	0.290121	2.469426	0.003
Display	**no of shared calls in clusters**	**-0.29593**	**-2.1196**	**0.017**	**-0.25063**	**-4.48097**	**<0.001**
Display	**no of shared clusters**	**-0.01068**	**-1.16169**	**0.114**	**-0.13759**	**-2.91584**	**0.002**
Display to the same female(s)	**local alignment**	**0.000141**	**0.585905**	**0.281**	**0.089959**	**3.385477**	**0.001**
Display to the same female(s)	**co-occurrence count**	**0.000873**	**9.093599**	**<0.001**	**0.171908**	**7.254592**	**<0.001**
Display to the same female(s)	**co-occurrence ratio**	**0.000483**	**6.049158**	**<0.001**	**0.123374**	**8.831332**	**<0.001**
Display to the same female(s)	**tri-gram**	**0.00053**	**8.094362**	**<0.001**	**0.157763**	**8.029798**	**<0.001**
Display to the same female(s)	**combined metric syntax**	**0.000499**	**9.501156**	**<0.001**	**0.266955**	**10.46692**	**<0.001**
Display to the same female(s)	within cluster distnace	-9.43E-07	4.827329	<0.001	-0.07004	4.935054	<0.001
Display to the same female(s)	overal spectral distance	-1.11E-06	3.430896	<0.001	0.061633	3.532583	<0.001
Display to the same female(s)	**no of shared calls in clusters**	**0.168394**	**0.550297**	**0.283**	**0.328675**	**3.917541**	**<0.001**
Display to the same female(s)	**no of shared clusters**	**0.01058**	**0.148682**	**0.444**	**0.200824**	**2.721565**	**0.005**

Aggression	local alignment	-9.98E-05	2.683707	0.004	-0.01241	1.801903	0.037
Aggression	co-occurrence count	-0.00069	-1.74748	0.043	-0.15718	-2.4818	0.007
Aggression	co-occurrence ratio	-0.00199	0.73561	0.234	-0.15537	-0.27514	0.391
Aggression	tri-gram	-0.00209	-2.99484	0.003	-0.23635	-4.22803	<0.001
Aggression	combined metric syntax	-0.00092	-0.09304	0.472	-0.14489	-1.56859	0.057
Aggression	within cluster distnace	-4.79E-06	-0.98132	0.153	-0.18738	-2.93563	0.002
Aggression	overal spectral distance	-6.80E-06	-0.71505	0.235	-0.08966	-0.51608	0.303
Aggression	no of shared calls in clusters	0.073423	-1.4574	0.081	0.202943	1.973985	0.025
Aggression	no of shared clusters	0.01978	0.805455	0.206	0.273949	2.817795	0.003
rank difference	**local alignment**	**-0.00773**	**-0.67044**	**0.256**	**-0.18557**	**-1.13667**	**0.136**
rank difference	**co-occurrence count**	**0.013823**	**1.407702**	**0.088**	**0.107795**	**1.503419**	**0.073**
rank difference	**co-occurrence ratio**	**0.00147**	**1.585165**	**0.071**	**0.029399**	**1.822724**	**0.05**
rank difference	**tri-gram**	**0.013162**	**1.452685**	**0.086**	**0.166888**	**1.946138**	**0.032**
rank difference	combined metric syntax	0.00252	1.613868	0.062	0.07696	1.886187	0.036
rank difference	within cluster distnace	1.01E-05	2.20606	0.017	0.046898	1.948113	0.023
rank difference	overal spectral distance	-1.85E-05	0.877541	0.147	-0.05839	0.597053	0.264
rank difference	no of shared calls in clusters	0.307566	-1.45431	0.077	0.007853	-0.92149	0.168
rank difference	no of shared clusters	-0.15066	-1.87413	0.033	-0.12703	-1.70084	0.052

***Supplementary table 1**: Summary of statistical results comparing an acoustic similarity or distance measure to a selected behavioral measure. The table includes the slope of the simple regression, Spearman's rank correlation coefficient (rho), their respective p-values, and Z-scores*

Bibliography

1. Abbassi, P., & Burley, N. T. (2012). Nice guys finish last: Same-sex sexual behavior and pairing success in male budgerigars. *Behavioral Ecology*, *23*(4), 775–782. https://doi.org/10.1093/beheco/ars030
2. Altmann, J. (1974). Observational Study of Behavior: Sampling Methods. *Behaviour*, *49*(3/4), 227–267.
3. Ang, T. Z., & Manica, A. (2010). Aggression, segregation and stability in a dominance hierarchy. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1686), 1337–1343. https://doi.org/10.1098/rspb.2009.1839
4. Armstrong, G., Rahman, G., Martino, C., McDonald, D., Gonzalez, A., Mishne, G., & Knight, R. (2022). Applications and Comparison of Dimensionality Reduction Methods for Microbiome Data. Frontiers in Bioinformatics, 2, 821861. https://doi.org/10.3389/fbinf.2022.821861
5. Aubin, T., & Jouventin, P. (1998). Cocktail–party effect in king penguin colonies. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1406), 1665–1673. https://doi.org/10.1098/rspb.1998.0486
6. Auersperg, A. M. I., & Von Bayern, A. M. P. (2019). Who's a clever bird — now? A brief history of parrot cognition. *Behaviour*, *156*(5–8), 391–407. https://doi.org/10.1163/1568539X-00003550
7. Balsby, T. J. S., Momberg, J. V., & Dabelsteen, T. (2012). Vocal Imitation in Parrots Allows Addressing of Specific Individuals in a Dynamic Communication Network. *PLOS ONE*, *7*(11), e49747. https://doi.org/10.1371/journal.pone.0049747

8. Balsby, T. J. S., & Scarl, J. C. (2008). Sex-specific responses to vocal convergence and divergence of contact calls in orange-fronted conures (Aratinga canicularis). *Proceedings of the Royal Society B: Biological Sciences*, *275*(1647), 2147–2154. https://doi.org/10.1098/rspb.2008.0517
9. Bartlett, P., & Slater, P. J. B. (1999). The effect of new recruits on the flock specific call of budgerigars (*Melopsittacus undulatus*). *Ethology Ecology & Evolution*, *11*(2), 139–147. https://doi.org/10.1080/08927014.1999.9522832
10. Berg, K. S., Delgado, S., Okawa, R., Beissinger, S. R., & Bradbury, J. W. (2011). Contact calls are used for individual mate recognition in free-ranging green-rumped parrotlets, Forpus passerinus. *Animal Behaviour*, *81*(1), 241–248. https://doi.org/10.1016/j.anbehav.2010.10.012
11. Bhat, A. S., Sane, V. A., Seshadri, K. S., & Krishnan, A. (2022). Behavioural context shapes vocal sequences in two anuran species with different repertoire sizes. *Animal Behaviour*, *184*, 111–129. https://doi.org/10.1016/j.anbehav.2021.12.004
12. Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: A comparison of three non-human primate species. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00390
13. Boughman, J. W. (1998). Vocal learning by greater spear–nosed bats. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1392), 227–233. https://doi.org/10.1098/rspb.1998.0286
14. Bradbury, J. W., & Balsby, T. J. S. (2016). The functions of vocal learning in parrots. *Behavioral Ecology and Sociobiology*, *70*(3), 293–312. https://doi.org/10.1007/s00265-016-2068-4
15. Briefer, E., Aubin, T., Lehongre, K., & Rybak, F. (2008). How to identify dear

enemies: The group signature in the complex song of the skylark *Alauda arvensis*. *Journal of Experimental Biology*, *211*(3), 317–326. https://doi.org/10.1242/jeb.013359

16. Briefer, E. F., Rybak, F., & Aubin, T. (2013). Does true syntax or simple auditory object support the role of skylark song dialect? *Animal Behaviour*, *86*(6), 1131–1137. https://doi.org/10.1016/j.anbehav.2013.09.019
17. Briseño-Jaramillo, M., Sosa-López, J. R., Ramos-Fernández, G., & Lemasson, A. (n.d.). *Flexible use of contact calls in a species with high fission–fusion dynamics*.
18. Brockway, B. F. (1964a). Ethological Studies of the Budgerigar (Melopsittacus Undulatus): Non-Reproductive Behavior. *Behaviour*, *22*(3–4), 193–222. https://doi.org/10.1163/156853964x00021
19. Brockway, B. F. (1964b). Ethological Studies of the Budgerigar: Reproductive Behavior. *Behaviour*, *23*(3–4), 294–323. https://doi.org/10.1163/156853964x00193
20. Brockway, B. F. (1965). Stimulation of ovarian development and egg laying by male courtship vocalization in budgerigars (Melopsittacus undulatus). Animal Behaviour, 13(4), 575–578. https://doi.org/10.1016/0003-3472(65)90123-5
21. Buhrman-Deever, S. C., Hobson, E. A., & Hobson, A. D. (2008). Individual recognition and selective response to contact calls in foraging brown-throated conures, Aratinga pertinax. *Animal Behaviour*, *76*(5), 1715–1725. https://doi.org/10.1016/j.anbehav.2008.08.007
22. Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012). Convergence and divergence in Diana monkey vocalizations. *Biology Letters*, *8*(3), 382–385. https://doi.org/10.1098/rsbl.2011.1182

23. Chase, I. D. (1974). Models of hierarchy formation in animal societies. *Behavioral Science*, *19*(6), 374–382. https://doi.org/10.1002/bs.3830190604
24. Cheney, D. L., Seyfarth, R. M., & Silk, J. B. (1995). The role of grunts in reconciling opponents and facilitating interactions among adult female baboons. *Animal Behaviour*, *50*(1), 249–257. https://doi.org/10.1006/anbe.1995.0237
25. Croft, D. P., Madden, J. R., Franks, D. W., & James, R. (2011). Hypothesis testing in animal social networks. *Trends in Ecology & Evolution*, *26*(10), 502–507. https://doi.org/10.1016/j.tree.2011.05.012
26. Dahlin, C. R., Smith-Vidaurre, G., Genes, M. K., & Wright, T. F. (2024). Widespread cultural change in declining populations of Amazon parrots. Proceedings of the Royal Society B: Biological Sciences, 291(2029), 20240659. https://doi.org/10.1098/rspb.2024.0659
27. Dahlin, C. R., Young, A. M., Cordier, B., Mundry, R., & Wright, T. F. (2014). A test of multiple hypotheses for the function of call sharing in female budgerigars, Melopsittacus undulatus. *Behavioral Ecology and Sociobiology*, *68*(1), 145–161. https://doi.org/10.1007/s00265-013-1631-5
28. Dale, J., Lank, D. B., & Reeve, H. K. (2001). Signaling Individual Identity versus Quality: A Model and Case Studies with Ruffs, Queleas, and House Finches. *The American Naturalist*, *158*(1), 75–86. https://doi.org/10.1086/320861
29. Dapporto, L., Santini, A., Dani, F. R., & Turillazzi, S. (2007). Workers of a Polistes Paper Wasp Detect the Presence of Their Queen by Chemical Cues. *Chemical Senses*, *32*(8), 795–802. https://doi.org/10.1093/chemse/bjm047
30. David, H. A. (n.d.). *Ranking from unbalanced paired-comparison data*.

31. De Vries, H. (1995). An improved test of linearity in dominance hierarchies containing unknown or tied relationships. *Animal Behaviour*, *50*(5), 1375–1389. https://doi.org/10.1016/0003-3472(95)80053-0
32. De Vries, H., Stevens, J. M. G., & Vervaecke, H. (2006). Measuring and testing the steepness of dominance hierarchies. *Animal Behaviour*, *71*(3), 585–592. https://doi.org/10.1016/j.anbehav.2005.05.015
33. Demartsev, V., Strandburg-Peshkin, A., Ruffner, M., & Manser, M. (2018). Vocal Turn-Taking in Meerkat Group Calling Sessions. *Current Biology*, *28*(22), 3661-3666.e3. https://doi.org/10.1016/j.cub.2018.09.065
34. Emory, G. R. (1976). Aspects of Attention, Orientation, and Status Hierarchy in Mandrills (Mandrillus Sphinx) and Gelada Baboons (Theropithecus Gelada). *Behaviour*, *59*(1–2), 70–87. https://doi.org/10.1163/156853976X00479
35. Ester Martin, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD'96). AAAI Press, 226–231.
36. Farabaugh, S. M., Brown, E. D., & Dooling, R. J. (1992). ANALYSIS OF WARBLE SONG OF THE BUDGERIGAR *MELOPSITTACUS UNDULATUS*. *Bioacoustics*, *4*(2), 111–130. https://doi.org/10.1080/09524622.1992.9753211
37. Farabaugh, S. M., Linzenbold, A., & Dooling, R. J. (1994). Vocal plasticity in budgerigars (Melopsittacus undulatus): Evidence for social factors in the learning of contact calls. *Journal of Comparative Psychology*, *108*(1), 81–92. https://doi.org/10.1037/0735-7036.108.1.81
38. Farine, D. R., & Carter, G. G. (2022). Permutation tests for hypothesis testing

with animal social network data: Problems and potential solutions. *Methods in Ecology and Evolution*, *13*(1), 144–156. https://doi.org/10.1111/2041-210X.13741

39. 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
40. Goodwin, S. E., & Podos, J. (2014). Team of rivals: Alliance formation in territorial songbirds is predicted by vocal signal structure. Biology Letters, 10(2), 20131083. https://doi.org/10.1098/rsbl.2013.1083
41. Girard-Buttoz, C., Bortolato, T., Laporte, M., Grampp, M., Zuberbühler, K., Wittig, R. M., & Crockford, C. (2022). Population-specific call order in chimpanzee greeting vocal sequences. *iScience*, *25*(9), 104851. https://doi.org/10.1016/j.isci.2022.104851
42. Guhl, A. M. (1956). The Social Order of Chickens. *Scientific American*, *194*(2), 42–45. https://doi.org/10.1038/scientificamerican0256-42
43. Hasegawa, T., & Soma, M. (2004). The Effect of Social Facilitation and Social Dominance on Foraging Success of Budgerigars in an Unfamiliar Environment. *Behaviour*, *141*(9), 1121–1134. https://doi.org/10.1163/1568539042664560
44. Hersh, T. A., Gero, S., Rendell, L., Cantor, M., Weilgart, L., Amano, M., Dawson, S. M., Slooten, E., Johnson, C. M., Kerr, I., Payne, R., Rogan, A., Antunes, R., Andrews, O., Ferguson, E. L., Hom-Weaver, C. A., Norris, T. F., Barkley, Y. M., Merkens, K. P., … Whitehead, H. (2022). Evidence from sperm whale clans of symbolic marking in non-human cultures. *Proceedings of the*

National Academy of Sciences, *119*(37), e2201692119. https://doi.org/10.1073/pnas.2201692119

45. Hile, A. G., Plummer, T. K., & Striedter, G. F. (2000). Male vocal imitation produces call convergence during pair bonding in budgerigars, Melopsittacus undulatus. *Animal Behaviour*, *59*(6), 1209–1218. https://doi.org/10.1006/anbe.1999.1438
46. Hile, A. G., & Striedter, G. F. (2000). Call Convergence within Groups of Female Budgerigars (Melopsittacus undulatus). *Ethology*, *106*(12), 1105–1114. https://doi.org/10.1046/j.1439-0310.2000.00637.x
47. Hobson, E. A., Mønster, D., & DeDeo, S. (2021). Aggression heuristics underlie animal dominance hierarchies and provide evidence of group-level social information. Proceedings of the National Academy of Sciences, 118(10), e2022912118. https://doi.org/10.1073/pnas.2022912118
48. Hobson, E. A. (2022). Quantifying the dynamics of nearly 100 years of dominance hierarchy research. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1845), 20200433. https://doi.org/10.1098/rstb.2020.0433
49. Hobson, E. A., Avery, M. L., & Wright, T. F. (2014). The socioecology of Monk Parakeets: Insights into parrot social complexity. *The Auk*, *131*(4), 756–775. https://doi.org/10.1642/AUK-14-14.1
50. Hobson, E. A., & DeDeo, S. (2015). Social Feedback and the Emergence of Rank in Animal Society. *PLOS Computational Biology*, *11*(9), e1004411. https://doi.org/10.1371/journal.pcbi.1004411
51. Hyman, J. (2003). Countersinging as a signal of aggression in a territorial songbird. *Animal Behaviour*, *65*(6), 1179–1185.

https://doi.org/10.1006/anbe.2003.2175

52. Jack, R. E., & Schyns, P. G. (2015). The Human Face as a Dynamic Tool for Social Communication. *Current Biology*, *25*(14), R621–R634. https://doi.org/10.1016/j.cub.2015.05.052

53. Janik, V. M., Sayigh, L. S., & Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences*, *103*(21), 8293–8297. https://doi.org/10.1073/pnas.0509918103

54. K. Lisa Yang Center for Conservation Bioacoustics. (2024). *Raven Pro: Interactive Sound Analysis Software (Version 1.6.5)* (Cornell Lab of Ornithology Patent). https://www.ravensoundsoftware.com

55. Kaprykowsky, H., & Rodet, X. (2006). Globally Optimal Short-Time Dynamic Time Warping, Application to Score to Audio Alignment. *2006 IEEE International Conference on Acoustics Speed and Signal Processing Proceedings*, *5*, V-249-V–252. https://doi.org/10.1109/ICASSP.2006.1661259

56. Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., Bohn, K., Cao, Y., Carter, G., Cäsar, C., Coen, M., DeRuiter, S. L., Doyle, L., Edelman, S., Ferrer-i-Cancho, R., Freeberg, T. M., Garland, E. C., Gustison, M., Harley, H. E., … Zamora-Gutierrez, V. (2016). Acoustic sequences in non-human animals: A tutorial review and prospectus: Acoustic sequences in animals. *Biological Reviews*, *91*(1), 13–52. https://doi.org/10.1111/brv.12160

57. Kershenbaum, A., Bowles, A. E., Freeberg, T. M., Jin, D. Z., Lameira, A. R., & Bohn, K. (2014). Animal vocal sequences: Not the Markov chains we thought they were. *Proceedings of the Royal Society B: Biological Sciences*,

281(1792), 20141370. https://doi.org/10.1098/rspb.2014.1370

58. Kershenbaum, A., & Garland, E. C. (2015). Quantifying similarity in animal vocal sequences: Which metric performs best? *Methods in Ecology and Evolution*, *6*(12), 1452–1461. https://doi.org/10.1111/2041-210X.12433
59. King, S. L., Harley, H. E., & Janik, V. M. (2014). The role of signature whistle matching in bottlenose dolphins, Tursiops truncatus. *Animal Behaviour*, *96*, 79–86. https://doi.org/10.1016/j.anbehav.2014.07.019
60. Kondo, N., & Hiraiwa-Hasegawa, M. (2015). The influence of social dominance on calling rate in the Large-billed Crow (Corvus macrorhynchos). *Journal of Ornithology*, *156*(3), 775–782. https://doi.org/10.1007/s10336-015-1191-8
61. Krams, I., Krama, T., Freeberg, T. M., Kullberg, C., & Lucas, J. R. (2012). Linking social complexity and vocal complexity: A parid perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1597), 1879–1891. https://doi.org/10.1098/rstb.2011.0222
62. Landau, H. G. (1951). On dominance relations and the structure of animal societies: I. Effect of inherent characteristics. *The Bulletin of Mathematical Biophysics*, *13*(1), 1–19. https://doi.org/10.1007/BF02478336
63. Lesk, A. M. (2014). Introduction to bioinformatics (Fourth edition). Oxford University Press.
64. Lehongre, K., Aubin, T., Robin, S., & Del Negro, C. (2008). Individual Signature in Canary Songs: Contribution of Multiple Levels of Song Structure. *Ethology*, *114*(5), 425–435. https://doi.org/10.1111/j.1439-0310.2008.01486.x
65. Levenshtein, V. I. (1966). Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady*, *10*, 707.

66. Maciej, P., Ndao, I., Hammerschmidt, K., & Fischer, J. (2013). Vocal communication in a complex multi-level society: Constrained acoustic structure and flexible call usage in Guinea baboons. *Frontiers in Zoology*, *10*(1), 58. https://doi.org/10.1186/1742-9994-10-58

67. Madabhushi, A. J., Bhat, A. S., & Krishnan, A. (2023). Allopatric montane wren-babblers exhibit similar song notes but divergent vocal sequences. Behavioral Ecology and Sociobiology, 77(10), 109. https://doi.org/10.1007/s00265-023-03385-9

68. Madabhushi, A. J., Wewhare, N., Binwal, P., Agarwal, V., & Krishnan, A. (2023). Higher-order dialectic variation and syntactic convergence in the complex warble song of budgerigars. *Journal of Experimental Biology*, jeb.245678. https://doi.org/10.1242/jeb.245678

69. Marino, L. (2017). Thinking chickens: A review of cognition, emotion, and behavior in the domestic chicken. *Animal Cognition*, *20*(2), 127–147. https://doi.org/10.1007/s10071-016-1064-4

70. Mathevon, N., Koralek, A., Weldele, M., Glickman, S. E., & Theunissen, F. E. (2010). What the hyena's laugh tells: Sex, age, dominance and individual signature in the giggling call of Crocuta crocuta. *BMC Ecology*, *10*(1), 9. https://doi.org/10.1186/1472-6785-10-9

71. McFee, B., Raffel, C., Liang, D., Ellis, D., McVicar, M., Battenberg, E., & Nieto, O. (2015). *librosa: Audio and Music Signal Analysis in Python*. 18–24. https://doi.org/10.25080/Majora-7b98e3ed-003

72. Medvin, M. B., Stoddard, P. K., & Beecher, M. D. (1993). Signals for parent-offspring recognition: A comparative analysis of the begging calls of cliff swallows and barn swallows. *Animal Behaviour*, *45*(5), 841–850.

https://doi.org/10.1006/anbe.1993.1105

73. Miller, P. J. O., & Bain, D. E. (2000). Within-pod variation in the sound production of a pod of killer whales, Orcinus orca. *Animal Behaviour*, *60*(5), 617–628. https://doi.org/10.1006/anbe.2000.1503

74. Mitani, J. C., & Gros-Louis, J. (1998). Chorusing and Call Convergence in Chimpanzees: Tests of Three Hypotheses. *Behaviour*, *135*(8/9), 1041–1064.

75. Møller, A. P. (1987). Variation in badge size in male house sparrows Passer domesticus: Evidence for status signalling. *Animal Behaviour*, *35*(6), 1637–1644. https://doi.org/10.1016/S0003-3472(87)80056-8

76. Moran, I. G., Loo, Y. Y., Louca, S., Young, N. B. A., Whibley, A., Withers, S. J., Salloum, P. M., Hall, M. L., Stanley, M. C., & Cain, K. E. (2024). Vocal convergence and social proximity shape the calls of the most basal Passeriformes, New Zealand Wrens. *Communications Biology*, *7*(1), 575. https://doi.org/10.1038/s42003-024-06253-y

77. Moravec, M. L., Striedter, G. F., & Burley, N. T. (2006). Assortative Pairing Based on Contact Call Similarity in Budgerigars, *Melopsittacus undulatus*. *Ethology*, *112*(11), 1108–1116. https://doi.org/10.1111/j.1439-0310.2006.01267.x

78. Mount, D. W. (2001). Bioinformatics: Sequence and genome analysis. Cold Spring Harbor Laboratory Press.

79. O'Connell-Rodwell, C. E., Berezin, J. L., Pignatelli, A., & Rodwell, T. C. (2024). The use of vocal coordination in male African elephant group departures: Evidence of active leadership and consensus. *PeerJ*, *12*, e17767. https://doi.org/10.7717/peerj.17767

80. Oliveira, C., Wahlberg, M., Silva, M. A., Johnson, M., Antunes, R.,

Wisniewska, D. M., Fais, A., Gonçalves, J., & Madsen, P. T. (2016). Sperm whale codas may encode individuality as well as clan identity. *The Journal of the Acoustical Society of America*, *139*(5), 2860–2869. https://doi.org/10.1121/1.4949478

81. Owens, D., & Owens, M. (1996). Social dominance and reproductive patterns in brown hyaenas,Hyaena brunnea, of the central Kalahari desert. *Animal Behaviour*, *51*(3), 535–551. https://doi.org/10.1006/anbe.1996.0058
82. Paoli, T., Palagi, E., & Tarli, S. M. B. (2006). Reevaluation of dominance hierarchy in bonobos (*Pan paniscus*). *American Journal of Physical Anthropology*, *130*(1), 116–122. https://doi.org/10.1002/ajpa.20345
83. Peckre, L., Kappeler, P. M., & Fichtel, C. (2019). Clarifying and expanding the social complexity hypothesis for communicative complexity. *Behavioral Ecology and Sociobiology*, *73*(1), 11. https://doi.org/10.1007/s00265-018-2605-4
84. Penndorf, J., Ewart, K. M., Klump, B. C., Martin, J. M., & Aplin, L. M. (2023). Social network analysis reveals context‐dependent kin relationships in wild sulphur‐crested cockatoos *Cacatua galerita*. *Journal of Animal Ecology*, *92*(1), 171–182. https://doi.org/10.1111/1365-2656.13839
85. Penndorf, J., Farine, D., Martin, J., & Aplin, L. (2024). *Long-term stability of dominance hierarchies in a wild parrot with fission-fusion dynamics*. https://doi.org/10.1101/2024.05.15.594429
86. Penndorf, J., Farine, D. R., Martin, J. M., & Aplin, L. M. (2023). *Parrot politics: Social decision-making in wild parrots relies on both individual recognition and intrinsic markers*. https://doi.org/10.1101/2023.10.02.560599
87. Phillip Bonacich. (1987, March). *Power and Centrality: A Family of Measures*.

American Journal of Sociology Vol. 92, No. 5. http://www.jstor.org/stable/2780000?origin=JSTOR-pdf

88. Prat, Y., Taub, M., & Yovel, Y. (2016). Everyday bat vocalizations contain information about emitter, addressee, context, and behavior. *Scientific Reports*, *6*(1), 39419. https://doi.org/10.1038/srep39419

89. Preuschoft, S., Schaik, C. P. V., Matsumura, S., Okamoto, K., Kuester, J., Paul, A., Colmenares, F., Hofer, H., & East, M. L. (2001). CHAPTER 5 Dominance and Communication: Conflict Management in Various Social Settings. In F. Aureli (Ed.), *Natural Conflict Resolution* (pp. 77–105). University of California Press. https://doi.org/10.1525/9780520924932-009

90. Podos, J., & Warren, P. S. (2007). The Evolution of Geographic Variation in Birdsong. In Advances in the Study of Behavior (Vol. 37, pp. 403–458). Elsevier. https://doi.org/10.1016/S0065-3454(07)37009-5

91. Rauber, R., Kranstauber, B., & Manser, M. B. (2020). Call order within vocal sequences of meerkats contains temporary contextual and individual information. *BMC Biology*, *18*(1), 119. https://doi.org/10.1186/s12915-020-00847-8

92. Revelle, W. (2007). *psych: Procedures for Psychological, Psychometric, and Personality Research* (p. 2.4.12) [Dataset]. https://doi.org/10.32614/CRAN.package.psych

93. Rushen, J. (1982). The peck orders of chickens: How do they develop and why are they linear? *Animal Behaviour*, *30*(4), 1129–1137. https://doi.org/10.1016/S0003-3472(82)80203-0

94. Ruch, H., Zürcher, Y., & Burkart, J. M. (2018). The function and mechanism of vocal accommodation in humans and other primates. Biological Reviews, 93(2), 996–1013. https://doi.org/10.1111/brv.12382

95. Sánchez-Tójar, A., Schroeder, J., & Farine, D. R. (2018). A practical guide for inferring reliable dominance hierarchies and estimating their uncertainty. *Journal of Animal Ecology*, *87*(3), 594–608. https://doi.org/10.1111/1365-2656.12776

96. Scarl, J. C., & Bradbury, J. W. (2009). Rapid vocal convergence in an Australian cockatoo, the galah Eolophus roseicapillus. *Animal Behaviour*, *77*(5), 1019–1026. https://doi.org/10.1016/j.anbehav.2008.11.024

97. Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2008). Overlapping and matching of codas in vocal interactions between sperm whales: Insights into communication function. *Animal Behaviour*, *76*(6), 1977–1988. https://doi.org/10.1016/j.anbehav.2008.07.032

98. Seibert, L. M., & Crowell-Davis, S. L. (2001). Gender effects on aggression, dominance rank, and af®liative behaviors in a ¯ock of captive adult cockatiels (Nymphicus hollandicus). *Applied Animal Behaviour Science*.

99. Shizuka, D., & McDonald, D. B. (2012). A social network perspective on measurements of dominance hierarchies. *Animal Behaviour*, *83*(4), 925–934. https://doi.org/10.1016/j.anbehav.2012.01.011

100. Shrout, P. E., & Fleiss, J. L. (1979). *Intraclass Correlations: Uses in Assessing Rater Reliability*. *Vol. 86*(No. 2), 420–428.

101. Smeele, S. Q., Senar, J. C., McElreath, M. B., & Aplin, L. M. (2024). *The effect of social structure on vocal flexibility in monk parakeets*. https://doi.org/10.1101/2024.09.20.614070

102. Smith, T. F., & Waterman, M. S. (1981). Identification of common molecular subsequences. *Journal of Molecular Biology*, *147*(1), 195–197. https://doi.org/10.1016/0022-2836(81)90087-5

103. Sugiura, H. (2007). Adjustment of Temporal Call Usage During Vocal Exchange of Coo Calls in Japanese Macaques. *Ethology*, *113*(6), 528–533. https://doi.org/10.1111/j.1439-0310.2007.01349.x

104. Thomsen, H. M., Balsby, T. J. S., & Dabelsteen, T. (2021). Follow the leader? Orange-fronted conures eavesdrop on conspecific vocal performance and utilise it in social decisions. *PLOS ONE*, *16*(6), e0252374. https://doi.org/10.1371/journal.pone.0252374

105. Tibbetts, E. A., & Dale, J. (2007). Individual recognition: It is good to be different. *Trends in Ecology & Evolution*, *22*(10), 529–537. https://doi.org/10.1016/j.tree.2007.09.001

106. Tibbetts, E. A., Pardo-Sanchez, J., & Weise, C. (2022). The establishment and maintenance of dominance hierarchies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1845), 20200450. https://doi.org/10.1098/rstb.2020.0450

107. Tu, H.-W., Osmanski, M. S., & Dooling, R. J. (2011). Learned vocalizations in budgerigars (*Melopsittacus undulatus*): The relationship between contact calls and warble song. *The Journal of the Acoustical Society of America*, *129*(4), 2289–2297. https://doi.org/10.1121/1.3557035

108. Tyack, P. L. (2020). A taxonomy for vocal learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1789), 20180406. https://doi.org/10.1098/rstb.2018.0406

109. Van Der Marel, A., Francis, X., O'Connell, C. L., Estien, C. O., Carminito,

C., Moore, V. D., Lormand, N., Kluever, B. M., & Hobson, E. A. (2023). Perturbations highlight importance of social history in parakeet rank dynamics. *Behavioral Ecology*, *34*(3), 457–467. https://doi.org/10.1093/beheco/arad015

110. Van Dyk, D. A., & Evans, C. S. (2008). Opponent assessment in lizards: Examining the effect of aggressive and submissive signals. *Behavioral Ecology*, *19*(4), 895–901. https://doi.org/10.1093/beheco/arn052

111. Van Lawick-Goodall, J. (1968). The Behaviour of Free-living Chimpanzees in the Gombe Stream Reserve. *Animal Behaviour Monographs*, *1*, 161-IN12. https://doi.org/10.1016/S0066-1856(68)80003-2

112. Vehrencamp, S. L., Ritter, A. F., Keever, M., & Bradbury, J. W. (2003). Responses to Playback of Local vs. Distant Contact Calls in the Orange-Fronted Conure, *Aratinga canicularis*. *Ethology*, *109*(1), 37–54. https://doi.org/10.1046/j.1439-0310.2003.00850.x

113. Wanker, R., Apcin, J., Jennerjahn, B., & Waibel, B. (1998). Discrimination of different social companions in spectacled parrotlets (Forpus conspicillatus): Evidence for individual vocal recognition. *Behavioral Ecology and Sociobiology*, *43*(3), 197–202. https://doi.org/10.1007/s002650050481

114. Wanker, R., Sugama, Y., & Prinage, S. (2005). Vocal labelling of family members in spectacled parrotlets, Forpus conspicillatus. *Animal Behaviour*, *70*(1), 111–118. https://doi.org/10.1016/j.anbehav.2004.09.022

115. Watts, D. P. (2018). Male dominance relationships in an extremely large chimpanzee community at Ngogo, Kibale National Park, Uganda. *Behaviour*, *155*(13–15), 969–1009. https://doi.org/10.1163/1568539X-00003517

116. Wewhare, N., & Krishnan, A. (2024). Individual-specific associations between warble song notes and body movements in budgerigar courtship

displays. *Biology Open*, bio.060497. https://doi.org/10.1242/bio.060497

117. Wittemyer, G., & Getz, W. M. (2007). Hierarchical dominance structure and social organization in African elephants, Loxodonta africana. *Animal Behaviour*, *73*(4), 671–681. https://doi.org/10.1016/j.anbehav.2006.10.008

118. Wright, T. F., & Dahlin, C. R. (2018). Vocal dialects in parrots: Patterns and processes of cultural evolution. *Emu - Austral Ornithology*, *118*(1), 50–66. https://doi.org/10.1080/01584197.2017.1379356

119. Wyndham, E. (1980). Diurnal Cycle, Behaviour and Social Organization of the Budgerigar *Melopsittacus Undulatus*. *Emu - Austral Ornithology*, *80*(1), 25–33. https://doi.org/10.1071/MU9800025

120. Wyndham, E. (1983). Movements and Breeding Seasons of the Budgerigar. Emu - Austral Ornithology, 82(sup1), 276–282. https://doi.org/10.1071/MU9820276s

121. Zhao, Z., Teoh, H. K., Carpenter, J., Nemon, F., Kardon, B., Cohen, I., & Goldberg, J. H. (2023). Anterior forebrain pathway in parrots is necessary for producing learned vocalizations with individual signatures. *Current Biology*, *33*(24), 5415-5426.e4. https://doi.org/10.1016/j.cub.2023.11.014

122. Zürcher, Y., Willems, E. P., & Burkart, J. M. (2021). Trade-offs between vocal accommodation and individual recognisability in common marmoset vocalizations. *Scientific Reports*, *11*(1), 15683. https://doi.org/10.1038/s41598-021-95101-8