Proceedings B. 10.1098/rspb.2018.1973
Supplementary Materials ESM 1

Dynamic network partnerships and social contagion drive cooperation

Roslyn Dakin and T. Brandt Ryder

MATERIAL AND METHODS

Proximity Data-Logging

Given that wire-tailed manakin social interactions take place in fixed locations (i.e., stable territories), individual variation in social behavior and population social structure can be observed by tracking male-male interactions at territories within each lek. Territorial and floater males were outfitted with coded nano-tags (NTQB-2, Lotek Wireless; 0.35g or ~3% of male body mass). These nano-tags transmit over a single VHF frequency (165 MHz), with each tag transmitting a unique digital ID. Tags were programmed with a 20 second pulse rate and a 12 hour on-off duty cycle for continuous monitoring over ~90 days. To measure manakin behavior, we used proximity data-loggers (hereafter DL; SRX-DL800, Lotek Wireless) with short whip antennas in the male territories. DLs were set with a gain range of 60 to 70 dB to constrain detections within each territory (~30 m radius, Fig. 2). Although the vast majority (> 95%) of receivers had gain set at 65 dB, the spatial proximity and/or vegetation density of certain territories occasionally required gain reduction (60 dB) or gain increase (70 dB) to maintain consistent detection radii. The resulting data stream recorded the ID of each unique tag detection, its date and time stamp, and a measure of the received signal strength in log decibel (dB) units (hereafter RSSI).

Behavioral Rule Set

Automated data-logging approaches can greatly increase the quantity and quality of data available to construct social networks and characterize behavior; however, these automated approaches require prior natural history knowledge of a particular system or behavior (Ryder *et al.* 2012). As such, we based our methods on previous studies of wire-tailed manakins showing that spatial co-occurrence within a territory is a both a necessary prerequisite for, and an accurate predictor of, long-term cooperative partnerships among males (Ryder *et al.* 2008, 2009, 2011, 2012). Thus, we defined a social interaction as two tags that were detected within 45 seconds of one another with a difference in RSSI values (Δ RSSI) < 10, indicating close spatial proximity (Fig. 2). This temporal threshold (< 45 s) was chosen to allow for the fact that each tag pings with a 20 second pulse rate, such that overlapping individuals could potentially have a 40 s gap between their respective pings. The spatial threshold was chosen to include only the pairs that were closest to each other within the detection radius. For a given male-male dyad, a second co-occurrence in the same location within 5 minutes was considered to be part of the same social interaction, but after a gap of $\geq$ 5 minutes, it was considered to be a new interaction between those two males.

Although this system was designed to minimize the detection radius, there were cases when DLs in adjacent territories had temporally overlapping detections of the same bird or dyad. These overlapping dyad detections could indicate a pair of birds that moved between two spatial locations, or they could indicate a pair that was located between two nearby DLs (e.g., in a position where the two birds could be detected by more than one DL at the same time). We consequently filtered the raw social interaction data (n = 40,127 interactions) using the following

procedure. First, we ordered the interaction data chronologically by start times, and identified any interactions involving the same dyad that had temporal overlap based on their start and end times. We then calculated an adjusted duration for each interaction by subtracting its overlapping time(s). The first interaction that failed to meet the criterion of having at least 45 seconds of non-overlapping duration was then removed from the data. We repeated this procedure iteratively, until only interactions with a non-overlapping duration of at least 45 seconds remained. In total, this left 36,885 distinct social interactions for further analysis.

Partner Dynamics

A prerequisite for testing indirect effects is that partnerships vary within focal individuals (Fig. 2f). Thus, we computed two descriptive statistics that capture (i) the daily turnover in partner identities, and (ii) the change in relative partner edge weights, respectively, in the manakin system. For our first statistic (partner identity turnover), we first calculated a measure of partner consistency, and then took the inverse of consistency as our measure of turnover. The steps for this calculation were as follows: for each focal bird in our analysis, we first calculated the proportion of days on which he interacted with each unique partner (considering only the top four partners per day, given this cut-off in our statistical analysis). We call this metric *C* for consistency and it can take values from $\frac{1}{d}$ to 1, where *d* is the number of days the focal bird had interactions. Then, we calculated an average *C* per day for the focal bird, took the grand mean, and normalized it on a scale from 0 to 1:

$$C_{focal} = \frac{\left(C_{avg} - \frac{1}{d}\right)}{\left(1 - \frac{1}{d}\right)}$$

Hence, C_{focal} captures how consistently a focal bird interacted with particular partners on a day-to-day basis. We took the inverse $(1 - C_{\text{focal}})$*100% as our measure of partner turnover. Thus, a turnover score of 0% represents complete consistency (no change whatsoever in a male's major partners), whereas a score of 100% represents complete turnover (a male has maximal partnership dynamics). Using this method, we determined that the manakins in our dataset exhibit considerable day-to-day turnover in the identities of their partners (mean = 53% ± SE 2%, n = 135 focal males with >1 day in the analysis).

Given that our variance-partitioning analysis accounts for partner weights, we also computed a second statistic to describe weight dynamics. To do this, we first expressed the top four partnerships per day as percentage weights within focal individuals (i.e., normalized within each focal bird). Then, for each focal bird, we computed the interval max% – min% for each of his partners. The average of this value captures the magnitude of his partner weight dynamics. We found that the manakins also exhibited substantial day-to-day variation in relative partner weights (mean = 45% ± SE 2%).

Ground-Truthing Experiment

To verify the spatial threshold for defining social interactions in our study system, we performed a ground-truthing experiment. Paired tags were moved sequentially along four 30 m-long transects centered on a receiver. The tags were briefly held stationary at 5 m intervals (0 m, 5 m, 10 m… 25 m, in each direction; n = 2,196 measurements total). Although RSSI values decay with distance, the resulting data shown in Fig. 2 illustrate how other factors such as topographic relief, vegetation density and variation in tag signal strength (±3 dB) can contribute to intra- and

inter-tag variation in RSSI. To validate our social interaction thresholds, we used these data to reconstruct all possible ΔRSSI values for two tags along the same transect with a known inter-tag distance (n = 1,216,897 pairs). The results are shown in Fig. 2c. When ΔRSSI was < 10, the tags were frequently in close spatial proximity (median inter-tag distance = 5 m), a distance at which two manakins in a display territory would be in visual and acoustic contact. In contrast, when ΔRSSI was ≥ 10, the tags tended to be much further apart (median distance = 15 m; distance > 5 m about 77% of the time). Together, these results confirm that our interaction data are conservative and include pairs that were most likely in close proximity. Previous studies have demonstrated that these physically close interactions are predictive of cooperative partnerships in this system (Ryder *et al.* 2008, 2009).

Behavioral Correlations

To describe correlations among the behavioral phenotypes, we fit a multivariate Gaussian model in MCMCglmm (Hadfield 2017) with the four behavioral phenotypes as dependent variables. To account for other sources of variation, this model included the same syntax described in Box 1 of the main text, except that partner identities were omitted. The key feature of this multivariate model was that it also included an unstructured covariance matrix for Focal.ID in the random effects and another unstructured covariance matrix for the residual (within-individual) variation (Houslay & Wilson 2017); this allowed us to compute the correlations among phenotypes while accounting for other sources of variation. We used the posterior variance and covariance estimates for each pair of phenotypes to calculate correlation coefficients, as follows:

$$Corr(PhenotypeA, PhenotypeB) = \frac{Cov(PhenotypeA, PhenotypeB)}{\sqrt{Var_{PhentoypeA} \times Var_{PhenotypeB}}}$$

This model was run with a burn-in of 3,000, storing every 100th sample for a total of 2,000 stored samples. The posterior correlation coefficients are shown in Table S2.

Null Models

To analyze the null (permuted) data, we fit MMMs using the same syntax presented in Box 1 of the main text. After verifying initial model diagnostics on a small number of model runs for each phenotype, we generated 1,000 null datasets, and fit a single model chain to each dataset storing 2,000 iterations. Because the direct and indirect effects are bounded by 0, we used one-sided tests to compare the observed posterior estimates with the median value obtained from these null permutations.

Finally, as an additional check on our results, we also examined null models that separately permuted either the Focal.ID labels or the Partner.ID labels alone. As expected, permuting Focal.ID reduced the direct effect to near zero, but gave the same indirect effect observed in the actual data, whereas permuting Partner.ID reduced the indirect effect to near zero, but gave the same direct effect observed in the actual results.

Sensitivity Analysis

Given the analytical limitations of how MMMs are implemented in current software, the covariance between phenotypic expression and social influence cannot be directly estimated; instead, we derived this parameter post-hoc as described in the main text. We therefore ran a sensitivity analysis to determine how accurately we could recover these parameters of interest.

We generated 200 simulated datasets with roughly equivalent direct and indirect effects (approx. 0.20 each) that had either positive, negative, or no covariance between the direct and indirect effect (phenotypic expression and social influence). The sample size in each simulated dataset was 2,000 repeated measures of 100 individuals, which is similar to, but less powerful than, our actual study (2,935 measures of 144 individuals). We then fit MMMs to each simulated dataset to recover the direct and indirect effects, and estimated the covariance as described in the main text. Fig. S2 shows the recovered results in relation to the simulated 'true' values. This analysis revealed that the covariance values are recovered with high fidelity at such a large sample size (Fig. S2a-b). Both the recovered direct effects and indirect effects were influenced by the magnitude of the true covariance, such that datasets with a relatively large-magnitude covariance tended to slightly underestimate the direct effect, and slightly overestimate the indirect effect (Fig. S2c-d). However, at n = 2,000, these biases were very small relative to the inherent uncertainty in the direct and indirect effect estimates. Specifically, for the direct effect, the bias (± 0.004 proportion units) is 25-fold smaller than the breadth of the 95% credible interval (approx. 0.10 units, similar to Fig. 4 and Table S3). For the indirect effect, the bias (± 0.037 units) is 4-fold smaller than the breadth of the 95% credible interval (approx. 0.15 units, also similar to Fig. 4 and Table S3). Given these ratios, the fact that we observe strong direct and indirect effects in our actual data (Fig. 4), and the greater statistical power of the real dataset, we feel confident that these small biases do not influence our conclusions.

Sample Sizes

A small number of individuals (n = 9) had tags that either stopped transmitting to the receivers shortly after they were released, or they left the study population, and hence they are not included in this study. Twenty-two of the successfully tagged males were also implanted with hormone and sham pellets in part of years 2 and/or 3 for a separate study of the endocrine mechanisms of cooperative behavior. We consequently limited our statistical analysis here to focal individuals at times when they had no implants and partners were known on the previous day; nevertheless, the implant birds still acted within the social networks and their interactions are retained when calculating the behavioral phenotype of other individuals.

References

Hadfield, J. (2017). *MCMCglmm 2.25: MCMC generalised linear mixed models*.

Houslay, T.M. & Wilson, A.J. (2017). Avoiding the misuse of BLUP in behavioural ecology. *Behav Ecol*, 28, 948–952.

Ryder, T.B., Blake, J.G., Parker, P.G. & Loiselle, B.A. (2011). The composition, stability, and kinship of reproductive coalitions in a lekking bird. *Behav Ecol*, 22, 282–290.

Ryder, T.B., Horton, B.M., Tillaart, M. van den, Morales, J.D.D. & Moore, I.T. (2012). Proximity data-loggers increase the quantity and quality of social network data. *Biology Letters*, 8, 917–920.

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

Ryder, T.B., Parker, P.G., Blake, J.G. & Loiselle, B.A. (2009). It takes two to tango: reproductive skew and social correlates of male mating success in a lek-breeding bird. *Proceedings of the Royal Society B*, 276, 2377–2384.

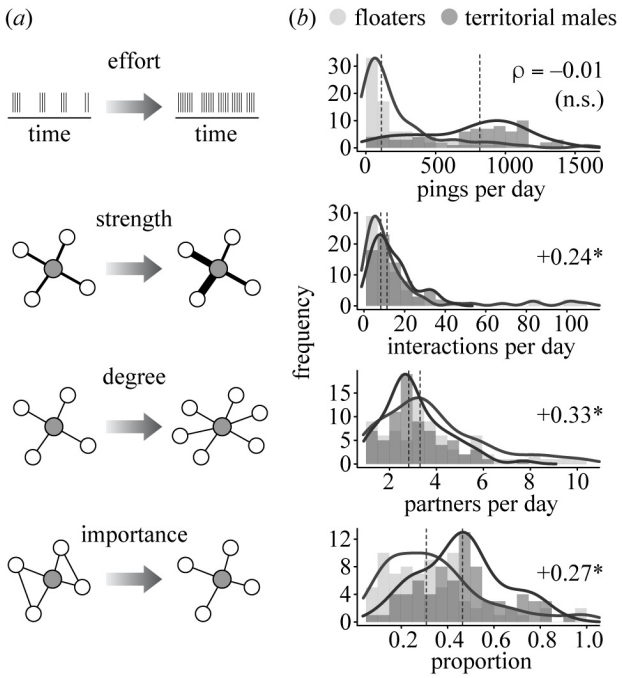

Fig. S1. We analyzed four behavioral phenotypes that are important for manakin reproductive success. (a) These behaviors occurred on territories within the leks and were quantified on a daily basis. The cartoon illustrates an increase in the phenotype from left to right. Note that in the illustration for importance, the focal male (gray node) maintains the same strength and degree, but increases his importance on the right, because each of his partners becomes more exclusive. Meanwhile, his partners' importance actually decreases on the right, demonstrating that changes in this metric are not necessarily positively correlated among partners (i.e., many scenarios are possible). (b) Observed distributions of the average phenotypic expression for floater and territory-holding males. Dashed vertical lines show the medians for each status class. For effort, 1,000 pings are approximately 5.5 hours of attendance. Strength, degree, and importance are all positively assorted (ρ) within the social network. See also Tables S1 and S2 for descriptive statistics.

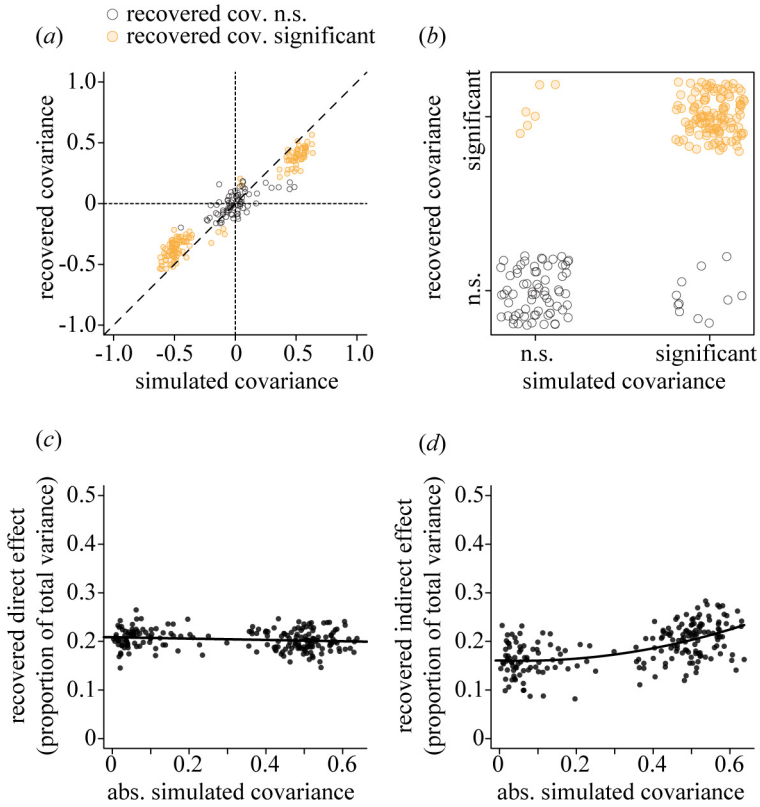

Fig. S2. Sensitivity analysis of direct and indirect effects. We simulated direct and indirect effects (proportion of total variance approx. 0.20 each), with either positive, negative, or no covariance, in a dataset of n = 2,000 measures of 100 individuals (vs. 2,935 measures of 144 individuals in the main text). A total of 200 simulations are shown. The "simulated" values are the true underlying effects and the "recovered" values are the estimates derived from the statistical analysis using the same methods described in the main text. (a) The recovered covariance (y-axis) closely tracks the true/simulated covariance (x-axis). The dashed line is the 1:1 line. (b) Error matrix for these covariance results (Type I error 6/200; Type II error 11/200). (c-d) Direct and indirect effects in relation to the magnitude of the simulated covariance. The change in the direct/indirect effects attributed to the magnitude of the covariance is much smaller than the uncertainty already inherent in those estimates.

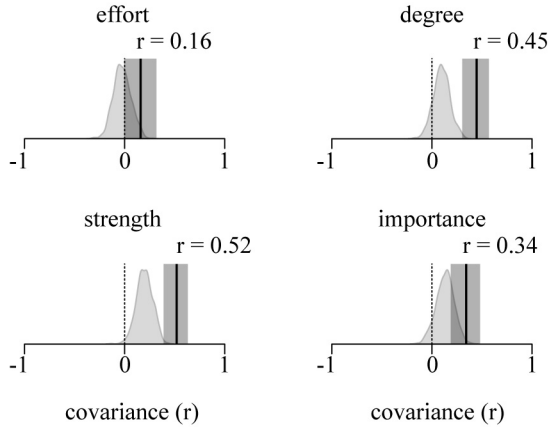

Fig. S3. Covariance between phenotypic expression and social influence, as compared to null expectations. The covariance between expression and social influence is estimated as the correlation coefficient (r) on a scale from –1 to 1; a value of 0 indicates no covariance. The observed covariance for each behavioral phenotype is shown as a black vertical line (± 95% CI shaded region). The light grey probability distribution gives the null expectation based on 1,000 permutations of the raw data. The observed covariance is significantly greater than 0 and greater than the null distribution for strength ($p < 0.001$), degree ($p < 0.001$), and importance ($p = 0.008$), indicating that these three phenotypes are more socially contagious that one would expect. Note that the null covariances for degree, strength, and importance also tended to be greater than 0, because the permutation preserved other spatiotemporal sources of covariance in the data, to provide a strong test.

Table S1. Descriptive statistics (grand mean ± SE) for n = 144 focal individuals in the MMM analysis (i.e., unmanipulated individuals with partnerships known on the previous day). Population means are calculated using the mean of individual means.

Phenotype	Year 1 (15-16)		Year 2 (16-17)		Year 3 (17-18)	
	Territorial n = 46	Floater n = 32	Territorial n = 41	Floater n = 48	Territorial n = 35	Floater n = 32
Effort (pings /day)	575 ± 59	101 ± 37	948 ± 59	321 ± 53	756 ± 61	149 ± 27
Strength (interactions /day)	8.2 ± 0.8	6.5 ± 1.0	26.2 ± 2.8	28.3 ± 5.0	8.3 ± 0.9	6.9 ± 0.8
Degree (partners /day)	2.6 ± 0.1	3.1 ± 0.3	4.5 ± 0.3	5.3 ± 0.4	2.3 ± 0.1	2.4 ± 0.2
Importance (proportion)	0.49 ± 0.03	0.35 ± 0.05	0.36 ± 0.03	0.22 ± 0.02	0.52 ± 0.03	0.49 ± 0.04

Table S2. Posterior estimates for the behavioral correlations among- and within-individuals. Correlation coefficients are on a scale from –1 to +1. Each cell gives the median posterior pairwise correlation estimate followed by the [95% central range] (n = 2,395 observations of 144 focal individuals used in the MMM analysis). This analysis follows the main variance-partitioning analysis by accounting for status, study year, and the individual's number of years of tenure in the study (fixed effects), as well as date and lek (random effects).

Among-individuals				
	Effort	Strength	Degree	Importance
Effort	--	0.66 *[0.52, 0.76]*	0.46 *[0.26, 0.60]*	0.15 *[–0.505, 0.35]*
Strength		--	0.87 *[0.81, 0.92]*	–0.14 *[–0.33, 0.07]*
Degree			--	–0.45 *[–0.61, –0.28]*
Importance				--

Within-individuals				
	Effort	Strength	Degree	Importance
Effort	--	0.47 *[0.44, 0.50]*	0.30 *[0.27, 0.34]*	0.15 *[0.11, 0.19]*
Strength		--	0.71 *[0.70, 0.73]*	0.10 *[0.06, 0.13]*
Degree			--	–0.22 *[–0.26, –0.19]*
Importance				--

Table S3. Results of the weighted MMM analysis of direct and indirect effects in manakin social networks. The fixed effects are estimated as the change in the response variable (phenotype) across status classes, study years, and the number of years of an individual's tenure in the study, respectively. For the random effects, the statistic is the proportion of total phenotypic variance (SD^2) explained on a scale from 0 to 1. R^2_{Bayes} is a measure of the variance explained by the whole model on a scale from 0 to 1. For the direct and indirect effects, statistical inference is based on a comparison with permuted null models. Additional tests for the other random effects (date and lek) are provided in Table S4.

Phenotype			Posterior median	95% Central range (lower, upper)	Comparison to null (p-value)
Effort	Fixed effects	Status (terr vs floa)	0.84	0.68, 0.99	
R^2_{Bayes} = 0.53		Study year (2 vs 1)	0.33	0.14, 0.53	
		Study year (3 vs 1)	0.12	–0.21, 0.46	
		Study tenure (# yrs)	0.15	–0.02, 0.32	
	Random effects	Direct effect	0.30	0.21, 0.38	0.003
		Indirect effect	0.12	0.07, 0.18	0.003
		Date	0.03	0.02, 0.05	
		Lek	0.09	0.03, 0.27	
		Residual	0.45	0.35, 0.52	
Strength	Fixed effects	Status (terr vs floa)	0.29	0.14, 0.44	
R^2_{Bayes} = 0.53		Study year (2 vs 1)	0.68	0.48, 0.87	
		Study year (3 vs 1)	–0.13	–0.47, 0.19	
		Study tenure (# yrs)	0.12	–0.05, 0.29	
	Random effects	Direct effect	0.24	0.17, 0.31	0.0003
		Indirect effect	0.19	0.13, 0.27	0.004
		Date	0.05	0.03, 0.08	
		Lek	0.09	0.02, 0.25	
		Residual	0.41	0.33, 0.48	
Degree	Fixed effects	Status (terr vs floa)	–0.02	–0.14, 0.12	
R^2_{Bayes} = 0.52		Study year (2 vs 1)	0.55	0.36, 0.73	
		Study year (3 vs 1)	–0.33	–0.62, –0.06	
		Study tenure (# yrs)	0.08	–0.06, 0.21	
	Random effects	Direct effect	0.12	0.08, 0.17	0.006
		Indirect effect	0.21	0.13, 0.30	0.003
		Date	0.08	0.06, 0.11	
		Lek	0.12	0.04, 0.31	
		Residual	0.46	0.35, 0.53	
Importance	Fixed effects	Status (terr vs floa)	0.26	0.12, 0.40	
R^2_{Bayes} = 0.49		Study year (2 vs 1)	–0.56	–0.75, –0.37	
		Study year (3 vs 1)	–0.03	–0.34, 0.28	
		Study tenure (# yrs)	0.23	0.07, 0.39	
	Random effects	Direct effect	0.17	0.12, 0.24	0.006
		Indirect effect	0.24	0.16, 0.33	0.005
		Date	0.03	0.02, 0.05	
		Lek	0.07	0.01, 0.25	
		Residual	0.46	0.37, 0.53	

Table S4. Model comparison for the spatiotemporal random effects, date and lek, using leave-one-out cross-validation. Smaller values of the leave-one-out information criterion (LOOIC) indicate a better model fit. The analysis compares the phenotype model with the full random effects structure from Box 1 to two other models where the random effect of date or lek had been removed. Delta is the change in LOOIC after removing the random effect.

Phenotype	Model	LOOIC (SE)	Delta
Effort	Full model	6475.24 (131.58)	0
	– Date	6542.09 (134.83)	66.85
	– Lek	6502.62 (130.86)	27.38
Strength	Full model	6484.25 (84.93)	0
	– Date	6654.17 (84.69)	169.92
	– Lek	6494.87 (84.91)	10.62
Degree	Full model	6554.99 (81.39)	0
	– Date	6806.05 (78.73)	251.06
	– Lek	6589.40 (81.01)	34.41
Importance	Full model	6739.38 (82.16)	0
	– Date	6811.44 (82.17)	72.07
	– Lek	6745.17 (82.42)	5.79