

Using playback to test leadership in mixed-species flocks and compare flocking with mobbing

Liping Zhou^{a, b, c}, Indika Peabotuwage^a, Kang Luo^{c, d}, Rui-Chang Quan^{b, *}, Eben Goodale^{a, *}

^a Guangxi Key Laboratory of Forest Ecology and Conservation, College of Forestry, Guangxi University, Nanning, Guangxi Zhuang Autonomous Region, China

^b Center for Integrative Conservation, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Science, Menglun, Mengla, Xishuangbanna, Yunnan, China

^c University of Chinese Academy of Sciences, Beijing, China

^d Ailaoshan Station for Subtropical Forest Ecosystem Studies, Chinese Academy of Sciences, Jingdong, Yunnan, China

ARTICLE INFO

Article history:

Received 9 July 2020

Initial acceptance 23 September 2020

Final acceptance 29 June 2021

MS. number: 20-00521R

Keywords:

animal sociality
bird vocalizations
group behaviour
keystone species
playback experiment
species interaction

Playback can help identify keystone species in communities that interact behaviourally. Specifically, playback has been used to identify leader species in mixed-species flocks (MSFs) of birds, although it is not clear whether responding heterospecifics are attracted to the leader species or the whole MSF. Playback can also simulate a mobbing response; both flocking and mobbing are regarded as adaptations to predation, but the species participating in these phenomena are not often compared. We performed a complex experiment, comparing heterospecific responses to playback of (1) MSFs (multiple species calling simultaneously), (2) a primary leader species, (3, 4) two secondary leader species that occasionally lead MSFs, (5, 6) two follower species, (7) one nonflocking species and (8) one predator (a small owl). In 240 trials over 2 years, 72 species made 980 heterospecific approaches to the speaker. We found the predator treatment attracted the most species, and the weakest response was to the non-flocking species. Other patterns were unexpected, however: the response to the MSF, and to the primary leader, was not higher than to other treatments that included MSF members, and the response to one follower species (the particularly vigilant grey-headed canary-flycatcher, *Culicicapa ceylonensis*) was higher than that to one secondary leader. Although flocking propensity did not influence the response, species that followed MSFs responded more than those that led them. There was a strong correlation between which bird species responded to the owl and which responded to any MSF member. Our results reveal (1) a response by flock followers to all MSF participants, as if they used these species' vocalizations to search for the whole MSF, (2) some weak evidence that vigilant species are particularly attractive, and (3) a general correspondence between the species that are attracted to an MSF and those that mob.

© 2021 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

Keystone species are found across a wide range of communities. Originally defined as top predators that regulate other species through trophic cascades (Paine, 1969), the concept of keystone species has more recently been broadened to include any species that has greater influence on other organisms than would be expected from the overall biomass of its population (Paine, 1995). Keystone species are usually thought of as ecologically interacting with other species; for example, certain microorganisms affect the whole community because of their capabilities in nutrient cycling

or decomposition (Banerjee et al., 2018) and humans affect community structure through their many disturbances (Worm & Paine, 2016). The effects of keystone species can be explicitly behavioural, in that some species termed 'community informants' may be particularly good information sources for other species (Hetrick & Sieving, 2012). The presence of such community informants can be an important factor organizing animal communities (Goodale et al., 2010), and in particular a mechanism of mutualism between members of a guild (Crowley & Cox, 2011).

In bird communities, a keystone role is played by the leaders of mixed-species flocks (MSFs). MSFs are found in many kinds of habitats (e.g. among pelagic, shorebird, waders, waterfowl, etc.), but they are most dominant in forested habitats, where in some

* Corresponding authors.

E-mail addresses: quanrc@xtbg.ac.cn (R. -C. Quan), ebengoodale@gxu.edu.cn (E. Goodale).

locations more than 50% of individuals can be found in MSFs at any one time (Eguchi et al., 1993; Latta & Wunderle, 1996). MSF participants tend to be insectivorous (Goodale et al., 2017), small (Sainz-Borgo et al., 2018) and vulnerable to predation (Thiollay, 2003). Leader species are those that have a tendency to be at the front or leading edge of MSFs (Zhou et al., 2019), and are a subset of a somewhat larger concept of 'nuclear' species, those that are particularly important to the formation or cohesion of MSFs (i.e., it is possible for a nuclear species not to be a leader, but to have some other function that brings birds into MSFs or keeps them there; Moynihan, 1962). Most reports of MSFs include some description of which species are nuclear or leaders (reviewed by Zou et al., 2018). Leader species tend to be gregarious (having many individuals per MSF), and are active and noisy (Goodale & Beauchamp, 2010; Hutto, 1994; Moynihan, 1962), although in some cases, they may not be particularly gregarious but be especially vigilant (Greig-Smith, 1981; Martínez et al., 2018). Leader species also tend to forage at a variety of vertical strata so that they can be followed by many other species (Hsieh & Chen, 2011). Because they have a large conspecific audience it has been suggested that these species may be information providers, and particularly provide costly information, such as alarm calls and calls of food recruitment, that may be useful to their conspecifics, but can also be eavesdropped on by other species (Farine et al., 2015; Goodale et al., 2010). Hence, it is argued that leader species can be considered community informants (Carlson et al., 2020; Hetrick & Sieving, 2012).

Playback can be used to examine questions about leadership in MSFs, as heterospecific attraction towards a species' vocalizations can be used as a measure of their desire to associate with that species (as pioneered by Mönkkönen et al., 1996). Leader species would be expected to attract the most heterospecific responses, as most species follow them in the field. Playback experiments can suggest what benefits follower species may be seeking in MSFs; for example, Goodale and Kotagama (2005b) asked whether follower species were more attracted to a 'sentinel' fly-catching leader, which provides information about predators, or a gregarious leaf-gleaning leader, which could provide foraging benefits as well as information about predation. However, a complication of the playback method is that birds could be attracted to a simulated leader not to accrue benefits from that particular species, but because that species might simply be an indication of the MSF as a whole (Goodale et al., 2020). To distinguish between benefits derived from the whole MSF and benefits derived from one leader species, one needs complex experiments that simulate the whole MSF system, as well as an array of leaders and followers. If birds are attracted to the group as a whole, they should be attracted to any species that is frequent in MSFs and has a high propensity to flock (i.e. most individuals are found in MSFs; Julien & Thiollay, 1998). In contrast, if birds are attracted to a particular leader species, there should be large differences in attraction between treatments, with the greatest attraction to the leader, and, if the leader is included in the MSF simulation, to the whole MSF treatment. An even more complex experiment would involve the simulation of 'secondary' leaders, species that lead MSFs occasionally or do so when primary leaders are absent (Diamond, 1987). If birds are attracted to particular leader species, then secondary leaders could be intermediately attractive, between primary leaders (those that normally lead MSFs) and followers. The response to the whole MSF treatment might not be as great as the sum of the responses to the single-species treatments, however, if secondary leaders are not attractive in the presence of primary leaders, or if birds are simply attracted to the flock as a whole.

Playback can also be a way to study another grouping phenomenon among birds: mobbing. As opposed to MSFs, which can usually be found throughout a day and are sometimes highly

stable across months or even, in an exceptional system where the whole MSF mutually defends a territory, decades (Martínez & Gomez, 2013), mobbing is a more temporary association. In mobbing, birds from the community gather around a stationary predator, acting aggressively towards the predator, to the point of driving it off, and often actively communicating (Carlson et al., 2017; Dutour et al., 2017a). Species that mob, especially towards small predators such as owlets, tend to be small (Dutour et al., 2017b). Mobbing can be elicited by using the playback of the predator's vocalizations and/or with a model (Hua & Sieving, 2016). This intriguing phenomenon has been the subject of intense scientific interest for a long time, with many explanations of why it is adaptive, including that it drives off the predator, that it does not allow the predator to make a surprise attack, that it can teach young about danger (Curio, 1978) and that it can even be a way for individuals to advertise their fitness (Dugatkin & Godin, 1992). As well as conspecifics, mobbing calls attract heterospecifics (Dutour et al., 2017a; Hurd, 1996), which from the caller's perspective might be particularly useful to dilute the risk of this dangerous situation without being either kin or mates. Interestingly, the initiators of mobbing assemblages often also seem to be leaders of MSFs, perhaps because they are preadapted to be vigilant about predators due to their large conspecific audiences (Jiang et al., 2020). Although these two phenomena, mixed-species flocking and mobbing, are generally considered to both be adaptations to reduce predation, the connections between them have rarely been probed, especially experimentally.

In this study, we compared heterospecific responses to simulations of many species in an MSF together, to a range of single species that vary in how often they lead MSFs and to an owlet playback call that stimulates mobbing. We conducted this experiment in the Ailaoshan forest of Yunnan Province, China, where mixed-species flocking is particularly frequent: almost 60% of individual birds are found in MSFs at any one time (Zhou et al., 2019). Our original experiment conducted in 2017 had seven treatments, including the MSF treatment, the primary leader, the Yunnan fulvetta, *Alcippe fratercula*, two secondary leaders, two follower species and a nonflocking control species. Heterospecific response to playback was weak, however, and the differences in responses between treatments were not very clear, so in 2018 we decided to add a stimulus known to have a strong effect, the playback of an owlet, as an extra treatment, providing a unique opportunity to compare responses to mobbing stimuli and to MSFs. As well as comparing the total number of species that responded in the treatments, we also compared the communities that responded to each treatment, in their multivariate composition (through ordination, and correlations between the vectors of responding species to each treatment), and in the traits of the species that compose them (the species' abundances, tendencies to participate in and follow MSFs, body sizes and diets). We hypothesized that MSF participants would be attracted to specific species that provide benefits to heterospecifics and hence are MSF leaders, more than to the whole MSF. Specifically, we predicted we would find the following patterns: (1) MSF-participating species that have a high propensity to flock (and hence are small and insectivorous), and particularly MSF-follower species (i.e. not including the leaders themselves) would come to the MSF treatment at the highest level, because it simulated the presence of several leader species; (2) MSF-participant/follower species would come to the individual species treatments in the following order: primary leader > secondary leaders > follower species > nonflocking species; (3) there would be some correspondence between the mobbing and the flocking treatments, because flock leaders tend to be mobbing initiators, and mobbers of owlets also tend to be small.

METHODS

Study Site

This experiment was conducted near the Ailaoshan Station for Subtropical Forest Ecosystem Studies, a research station of the Chinese Academy of Sciences. The research station is located inside the Ailaoshan National Nature Reserve (24°31' N, 101°01' E), Yunnan Province, China. The forest that surrounds the station is mid-montane moist evergreen broad-leaved forest (Zhu, 2016), at elevations ranging from 2300 to 2600 m above sea level; mean annual precipitation is ca. 1850 mm, with ca. 85% coming in the wet season, between May and October (Chen et al., 2019). We worked in primary forest, although some transects were on roads through primary forests that had been disturbed by the construction of the road.

Survey on MSFs and Species Abundance

The first step of our survey was collecting data on the composition of MSFs at the study site, the species that led these MSFs and the abundances of different species in this location (as part of an earlier project summarized by Zhou et al., 2019). We first placed three 1 km transects on infrequently travelled dirt roads in the forest, so that each transect was at least 500 m from another, following the recommendations on MSF surveys of Goodale et al. (2009). At least one observer and one note-taker walked these transects twice a day, once in the morning (between 0730 and 1100 hours), and once in the afternoon (between 1500 and 1800 hours), at a rate of 1 km per 1.5 h. In total, each transect was walked six times between October 2015 and January 2017 in the nonbreeding season (August–March, with the majority of observations between September and January), which is the season of greatest MSF activity. While walking a transect, we noted every bird we saw or heard and its estimated distance from the transect (i.e. a variable radius transect method), and recorded whether it was solitary, in a conspecific group or in an MSF. For the purposes of this study, we only used bird detections estimated to be equal to or less than 30 m from the transect (i.e. a belt transect method), judging that there were too few data points per species to use distance sampling methods to adjust for differences between species in detectability

(Buckland et al., 2001), and that we could assume 100% detection within 30 m. An MSF was defined as birds of two or more different species moving in the same direction for at least 5 min. When we encountered an MSF, we spent a minimum of 5 min and a maximum of 15 min determining the number of individuals of all species (its composition). The MSF in the area averaged 4.8 ± 2.8 (SD) species and 22.8 ± 14.5 individuals.

We also took data on the leadership of MSFs (Zhou et al., 2019). We noted every instance where we observed a bird clearly following a heterospecific (more exactly, if a second bird flew more than 3 m in the same direction as the first within 5 s) and recorded the species identities of the leader and follower. These data were collected in Ailaoshan and 230–280 km away in the Gaoligangshan National Natural Reserve, where most of the same species are present and the patterns of leadership appeared similar; we therefore pooled the data (see Supplementary Material).

Based on this information, we selected species that played different roles in MSFs for simulation by playback. Our criteria included that a species must be fairly abundant in Ailaoshan (>0.05 individuals per ha) and that it must be vocal with relatively loud vocalizations (>45 dB measured at 5 m; otherwise, playback would be futile). Selections included the following species: the primary leader, the species that led the most MSFs, Yunnan fulvetta (YNFU hereafter); secondary leaders, species that were sometimes leaders even in the presence of the primary leader YNFU or in its absence, chestnut-tailed minla, *Actinodura strigula* (CTMI) and red-tailed minla, *Minla ignotincta* (RTMI); followers, species that were often in an MSF but only led occasionally, grey-headed canary-flycatcher, *Culicicapa ceylonensis* (GHCF) and green-backed tit, *Parus monticolus* (GBTI); and a nonflocking species that was never seen in MSFs, the chestnut-headed tesia, *Cettia castaneocoronata* (CHTE). Data on these species' abundance and leadership in MSFs are presented in Table 1.

Recordings

For making recordings, we set up seven transects that were 1 km long on foot trails or dirt roads in the same forest (including the three that had been used in the flock survey before), with each transect at least 500 m from all others at all points. Recordings were made between 18 September 2017 and 23 October 2017. For

Table 1
The species that were selected to be simulated in the playback experiment

Species	Category	Species trait data					MSF following data			
		Body length (cm)	Abundance (individuals/ha)	% of MSFs in which seen (60 MSFs in total)	Flocking propensity (% individuals in MSF)	No. of individuals per MSF (\pm SD)	No. of MSFs in which species was present as a leader	No. of times this species was seen to lead	No. of MSFs in which species was present as a follower	No. of times this species was seen to follow
YNFU	Primary leader	12.5–14	1.58	23.3	90.0	11.7 ± 6.0	15	62	5	8
CTMI	Secondary leader	16–18.5	1.44	33.3	80.0	6.2 ± 3.7	12	38	6	16
RTMI	Secondary leader	13–14.5	0.08	6.7	100.0	2.3 ± 0.5	7	21	4	7
GHCF	Follower	12–13	0.07	5.0	50.0	1.3 ± 0.6	0	0	6	8
GBTI	Follower	12.5–13	0.31	35.0	82.4	1.4 ± 0.4	12	17	17	22
CHTE	Nonflocking	8–9.5	0.11	0.0	0.0	0	0	0	0	0
All	NA	NA	16.7	NA	60.8	NA	NA	NA	NA	NA

All simulated species with some data as to their abundance, their frequency in and their propensity to participate in mixed-species flocks (MSFs), and the amount they led MSFs or other individuals. Body length data from Birds of the World, on-line edition (<https://birdsoftheworld.org/bow/home>). Other species traits data were taken only from Ailaoshan Natural Nature Reserve, in the 2400–2500 elevation transects used for the playbacks. However, the MSF following data was also taken in Gaoligong Natural Nature Reserve (the data are separated between sites in the Supplementary Material). YNFU = Yunnan fulvetta, *Alcippe fratercula*, CTMI = chestnut-tailed minla, *Actinodura strigula*, RTMI = red-tailed minla, *Minla ignotincta*, GHCF = grey-headed canary-flycatcher, *Culicicapa ceylonensis*, GBTI = green-backed tit, *Parus monticolus*, CHTE = chestnut-headed tesia, *Cettia castaneocoronata*. NA = not applicable.

recording equipment, we used an omnidirectional microphone (Sennheiser ME62) embedded in a Telinga parabola and attached to a Marantz PMD 670 digital recorder. We walked these transects in the mornings (0830–1130 hours) and the afternoons (1500–1800 hours), avoiding early mornings because MSFs were not fully formed and some species have distinct morning chorus call types. Concentrating on the species selected to be simulated, we recorded both inside and outside MSFs, and we aimed to record each species on as many transects as possible. We noted when birds appeared to be alarmed (including to the observers) and avoided using these recordings.

During the recording period, we also measured the amplitude of vocalizations of the simulated species. For these measurements we found birds that were 5 m away from us in a straight line that were vocalizing (as long as they did not appear to be alarmed by our presence). We used a CEM DT-850 sound pressure level (SPL) meter, with A weighting and fast meter settings, to make measurements, and collected at least five measurements for each of the simulated species (Table A1).

Constructing Playback Tapes

From the recordings, we made five 1 min playback exemplars per simulated species. To make these exemplars we tried to sample widely from the recordings, so that for each treatment the different exemplars came from different transects. The recordings were selected to avoid alarm calls but were otherwise a random selection of vocalizations of the species. We used Raven 3.0 (Cornell Laboratory of Ornithology, Ithaca, NY, U.S.A.) software to make the playback tapes. For the six simulated species, recordings were constructed so that only these species were represented, and all other species' vocalizations and background noise were removed. Recordings were high-pass filtered to remove sounds below 500–700 Hz. We also adjusted the waveform so that amplitude peaks of the species were roughly similar across the 1 min recording. We then repeated each 1 min tape to obtain a 10 min tape.

In addition to the six simulated species, we made two other playback treatments. One was an MSF treatment, made to simulate the whole MSF. In other playback experiments, researchers have artificially constructed playback tapes that represented combinations of species, combining different species' vocalizations sequentially (Gu et al., 2017). In this experiment, however, many species were simulated separately, and serial combinations sounded artificial, so we opted for natural recordings of many species. In selecting these tapes, we specifically found recordings in which there were multiple species vocalizing together during a 1 min period (again, avoiding alarm calls). We required that the recording contained at least one of the simulated species and more than two species' vocalizations; the tapes contained on average the vocalizations of 6.4 ± 0.5 species (Table A2). A final treatment was the collared owl, *Glaucidium brodiei* (COOW), which was a common predator at the site, although difficult to record there. For playback of this species, we downloaded five COOW recordings from Xeno-Canto (<https://www.xeno-canto.org/>), which were all recorded in China and of the best signal-to-noise ratio ('A' quality). Again, from this material we made five exemplars of a 1 min playback tape, repeated to be 10 min in total.

Playback

For playback, we chose 15 locations from our recording transects (the beginning, middle and end of five transects), with each location at least 500 m apart from other ones. To reduce temporal variation, a location was always visited at a certain time of day (early morning; mid-to-late morning; afternoon), with all eight treatments played at

each site during one round of playback. In a round we used all 40 playback exemplars (8 treatments \times 5 exemplars), with each exemplar randomly assigned to one location. The order of playback (which treatment went first) was systematically varied among locations, and the time-of-day categories balanced, so that a treatment was evenly spread as to the time of day it was conducted. Each round of playback used five different locations, so that all 15 locations were used in three rounds; locations used on the same round were on different transects. We conducted three rounds of playback between 21 October and 22 November 2017 (after the recordings were made), and then another three rounds of playback between 27 September and 30 October 2018 (a time selected to be a little earlier seasonally than the previous year's playback, to check whether responses during late September and early October, the period in which the recordings were made, would be similar to responses later in the year). As explained above, the predator COOW treatment was not part of the original experiment, and both rounds of this treatment were in 2018.

Playbacks were broadcast using JTS WA-35 System speakers. Two playback speakers were placed ca. 5 m apart on either side of the centre of the location, tied to trees approximately 1.5 m off the ground, with the speakers facing opposite directions. We used the sound level measurements as a guide to the volume of the playback. Since the speakers are very directional, we played them ca. 6 dB louder than the field measurement of the SPL meter (to increase the chance they would be heard by birds outside the direction in which they projected). For the MSF treatment, we set the speaker for the correct volume for the loudest species. The volume of the predator COOW treatment was adjusted to be similar to natural amplitudes of this species by ear.

During a trial, two observers were positioned on opposite sides of the centre of the playback arena. Before starting a trial, we conducted a 7.5 min point count, noting the species that were present within 15 m of the centre of the arena, within 30 m, and then greater than 30 m. We did not conduct a trial if an MSF was detected in the area (at any distance), because multiple birds would come together. Playback was then conducted continuously for 10 min. During the playback period, the observers noted all the species seen or heard to approach the speaker, when they did so (relative to the start of playback) and the closest distance they got to the centre of the arena. Playback was then followed by a second 7.5 min point count. After the trial was over, the two observers met to compare their notes and make a final data sheet summarizing all bird responses, including a map of the playback arena and the birds' movements, which we used to judge whether species' responses were independent (see analysis below). Playback trials were not conducted in the rain, but were conducted in fog/mist, because misty conditions were prevalent.

Ethical Note

This project was approved by the Ailaoshan National Nature Reserve and by the Ailoshan Research Station of the Chinese Academy of Sciences. It followed the laws and regulations of the People's Republic of China and the ASAB/ABS Guidelines for the Use of Animals. A playback trial lasted only 10 min, and by the end of this time most respondents had dispersed. Trials were also spread out across 15 playback locations, and over 2 years, so we believe the impact on the birds' fitness was minimal.

Data Preparation

As we were interested in heterospecific responses, we first removed conspecific responses (i.e. the same species responding as was simulated) from the data (in this step we removed 55

responses). Because there were so many species simulated in the MSF treatment, we did not remove conspecifics' responses in the analysis of this treatment. Second, we were concerned that some species might have followed other species towards the playback location, and thus not represent independent responses to the playback. We defined a nonindependent response as one in which the bird came from the same direction as a bird that had responded earlier, and within 5 s. Using the maps of the birds' movements, we then removed 87 nonindependent responses from the data set.

Comparison of the Number of Species that Responded

All analyses were conducted using R version 3.6.2 (R Core Team, 2019). Although we considered $P < 0.05$ to be significant, we also report P values between 0.05 and 0.10. Mean values are shown \pm SD.

We used generalized linear mixed models (GLMM) with the 'lme4' package to analyse the number of species that responded in the trials of different treatments. The number of species that responded to a trial was the response variable in the model, treatment (with eight levels representing the different simulated sounds) was a fixed effect, and playback location was included as a random effect. We attempted to include playback exemplar as a second random effect, but it explained zero variance leading to a singularity problem, and so we removed it. Since responding conspecifics produced sound, and therefore may have amplified the treatment, we included conspecific responses (the number of individuals) as another fixed effect. However, we did not use the number of heterospecifics in the models, because so many different species responded and they varied greatly in their behaviours (some very vocal, others not at all).

We were unable to include the round of playback in the full model because of the differences between the owl and the other treatments in when they were conducted; however, in supplemental analyses, we investigated the influence of round, and the interaction between treatment and round, on (1) 2018 data only and (2) both years' data, excluding the owl treatment. Another supplemental analysis (3) looked at the effect of year (2017 versus 2018). We simplified models sequentially, removing nonsignificant interactions ($P > 0.05$) and then nonsignificant variables. Because the response variable was a count, we used a Poisson distribution with the link function of 'log'. Multiple comparisons used the 'multcomp' package and the Tukey HSD method to control the family-wide alpha value. Model fit for the GLMM was assessed following Nakagawa and Schielzeth (2013).

Community Composition of Respondents

To compare the species composition of responding birds in the different treatments (again only including heterospecific independent responses), we made a list ('vector') of each responding species and the number of times it responded to a particular treatment. We then looked at the correlations between the vectors for each pair of treatments; a correlation of 1 would represent the same list of responding species, with the same number of responses by each species. Because the data were not normal, we used Spearman correlations, and tested the significance through Spearman correlation tests.

To visualize the differences in species composition in the different treatments, we conducted a nonmetric multidimensional scaling (NMDS) plot, using the 'vegan' package. For this analysis, we considered all individuals of all species that responded during the two trials of a treatment at one location to be a community, and the NMDS was run on an abundance matrix. Hence, each treatment had 15 data points. We then used a Kruskal–Wallis ANOVA, followed by

pairwise Mann–Whitney tests, Bonferroni corrected, to ask whether NMDS scores (for both axes separately) of each treatment were statistically significant from each other or had similar variances. We later repeated the NMDS with only 2018 data, to see whether the year (and the fact that the COOW treatment was only conducted in 1 year) influenced the analysis (see [Appendix](#)).

Comparison of Responding Species and Their Traits

Another way to look at the response to playback is at the level of the responding species, and specifically to identify those species that responded nonrandomly to playback (i.e. responding more to some treatments than others). To do this, we ran a Fisher's exact test on a 2 (species responded versus species did not respond) \times 8 (treatment) frequency table, using the 'fisher.test' function in base R, for each responding species.

We also investigated the traits of the species that responded. First, we looked at abundance, because differences between species in abundance is a sampling bias: more abundant species have the opportunity to respond more. We used the data from the MSF survey, using all detections within 30 m, to estimate the number of individuals per 100 ha per species. We had 54 species for which there were abundance data (see [Appendix](#)). We then ran a linear model, with the total number of responses of the species in all trials of all treatments as the response variable and abundance per 100 ha as the explanatory variable. To better fit parametric assumptions, we log transformed both response and explanatory variables, after the addition of 1, so that the values were always greater than 1.

The next trait we considered was flocking propensity, defined as the percentage of individuals that were seen in MSFs (Jullien & Thiollay, 1998). Again, flocking propensity was taken from the MSF survey data ($N = 54$ species). For this analysis, we did not include responses during trials of the predator COOW or non-flocking CHTE treatments, as we did not expect flocking species to come to nonflocking CHTE treatments, and we expected all species to come to the predator COOW treatment, regardless of their propensity to flock. Also, knowing from the abundance analysis that abundance was an important predictor of response, we used as a response variable the number of responses adjusted for (i.e. divided by) abundance and log transformed (with the addition of 1). Here, a linear model asked whether flocking propensity predicted abundance-adjusted response.

The third trait we considered was an index of following versus leading MSFs. Here we used the individual level following data to calculate a 'following index' as the percentage of individuals observed for the species that were following other birds. For this analysis, we used only species that had at least three observations for the individual level leadership data ($N = 31$); again, the analysis excluded predator COOW and nonflocking CHTE treatments. The linear model asked whether the following index predicted abundance-adjusted response (log transformed, with the addition of 1).

We also explored whether the different treatments had responding birds of different body size or diet. Each individual bird that responded to a treatment was a data point and was assigned that species' body length (cm) or that species' diet. Body length and diet data were extracted from Birds of the World, on-line edition (<https://birdsoftheworld.org/bow/home>), with species classified into insectivores, frugivores and nectarivores if the description of the species indicated that most of the diet was invertebrate, fruit or nectar, respectively. Species that combined diets were classified as omnivores. We then made bar charts showing the average body length of respondents or what percentage of the respondents were of the different diet categories. The minor differences between

treatments in these graphs meant that statistical analyses were not necessary.

RESULTS

Comparison of the Number of Species that Responded

We recorded 72 species (including two unidentified species) that responded, making a total of 980 independent heterospecific responses. The cumulative number of species responding to each treatment ranged from a high of 43 to the predator COOW treatment to a low of 30 to the nonflocking CHTE treatment (Fig. A1). There were also 55 conspecific responses; because conspecific response was never a significant independent variable (i.e. it did not amplify the treatment), we removed it from the simplified models.

The number of respondent species per trial was significantly higher to the predator COOW treatment than to all the other treatments (Fig. 1; all $Z \geq 4.21$, all $P < 0.001$; overall $R^2_m = 0.26$, $R^2_c = 0.33$). At the other extreme, the nonflocking CHTE treatment had on average the lowest response, significantly lower than four treatments (all $Z \leq -3.74$, all $P < 0.005$), but not significantly lower than the follower GBTI ($Z = -1.76$, $P = 0.64$) and RTMI ($Z = -1.22$, $P = 0.92$) treatments. The order of response did not coincide, however, with the hypotheses (MSF > primary leader > secondary leader > follower). For example, the secondary leader RTMI treatment had significantly less response than the follower GHCF treatment.

Supplemental analyses showed that there was never an interaction between treatment and round or year and round. That is, the seasonality (three rounds per year) or year of the playback trials did not influence the way in which treatment affected the response (independent heterospecific responses). However, in analyses that included both years but not the predator COOW treatment, round and year were significant, with response being higher in 2017 than 2018, and specifically the second round tending ($P = 0.06$) to have more response than the sixth round (see Appendix).

Community Composition of Respondents

All correlations between vectors (of the number of responses per species) for all pairs of treatments were positive, and all

Table 2

Correlations between pairs of treatments in the numbers of individual responses per species

	COOW	MSF	YNFU	CTMI	RTMI	GBTI	GHCF	CHTE
COOW	1							
MSF	0.60	1						
YNFU	0.69	0.64	1					
CTMI	0.51	0.60	0.53	1				
RTMI	0.51	0.53	0.49	0.60	1			
GBTI	0.56	0.51	0.52	0.47	0.62	1		
GHCF	0.62	0.59	0.67	0.59	0.53	0.49	1	
CHTE	0.46	0.38	0.50	0.44	0.55	0.57	0.48	1

The numbers represent Spearman correlation coefficients. All correlations were significant at $P < 0.001$. Correlations greater than 0.65 are in bold. COOW = collared owl, *Glaucidium brodiei*, MSF = mixed-species flock, YNFU = Yunnan fulvetta, *Alcippe fratercula*, CTMI = chestnut-tailed minla, *Actinodura strigula*, RTMI = red-tailed minla, *Minla ignotincta*, GHCF = grey-headed canary-flycatcher, *Culicicapa ceylonensis*, GBTI = green-backed tit, *Parus monticolus*, CHTE = chestnut-headed tesia, *Cettia castaneocoronata*.

correlations were significant at $P < 0.001$ (Table 2). Two Spearman correlations were above 0.65: they were between the responses to the predator COOW treatment and the primary leader YNFU treatment and between YNFU and the follower GHCF treatment. The weakest correlation (0.38) was between the whole MSF treatment and the nonflocking CHTE treatment.

The NMDS illustrated that similar communities responded to the different treatments (Fig. 2). All the ellipses, representing 1 SD from the centroid of each treatment, overlapped. However, there were significant differences between treatments in their variance: the nonflocking CHTE treatment had significantly more variance on axis 1 than the predator COOW treatment (Mann–Whitney U test, Bonferroni adjusted for seven tests; $P < 0.001$), as well as the follower GBTI ($P = 0.001$), secondary leader CTMI ($P = 0.012$) and follower GHCF ($P = 0.019$) treatments. The nonflocking CHTE treatment also had more variance on axis 2 than the predator COOW ($P = 0.014$) and the whole MSF ($P = 0.040$) treatments. The low variance to the predator COOW treatment could have been caused by its presentation having less temporal variation (as it was conducted in only 1 year). Therefore, we produced another NMDS only for 2018. This NMDS showed qualitatively similar results, with the predator COOW treatment again having little variance, similar to the whole MSF treatment (see Appendix).

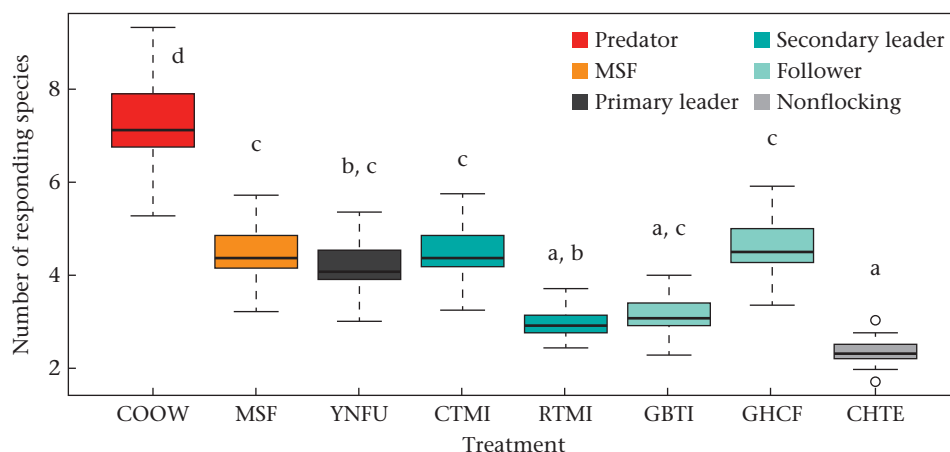


Figure 1. Number of species that responded per trial to the different treatments. The treatments are colour coded according to type (predator, MSF, primary leader, secondary leader, follower, nonflocking). Treatments with the same letter were not significantly different from each other. Box plots show the median value (middle line), the 25th and 75th percentiles (bottom and top of the box), with whiskers representing approximately 5% and 95% percentiles, and circles indicating any outlying points. COOW = collared owl, *Glaucidium brodiei*, MSF = mixed-species flock, YNFU = Yunnan fulvetta, *Alcippe fratercula*, CTMI = chestnut-tailed minla, *Actinodura strigula*, RTMI = red-tailed minla, *Minla ignotincta*, GHCF = grey-headed canary-flycatcher, *Culicicapa ceylonensis*, GBTI = green-backed tit, *Parus monticolus*, CHTE = chestnut-headed tesia, *Cettia castaneocoronata*.

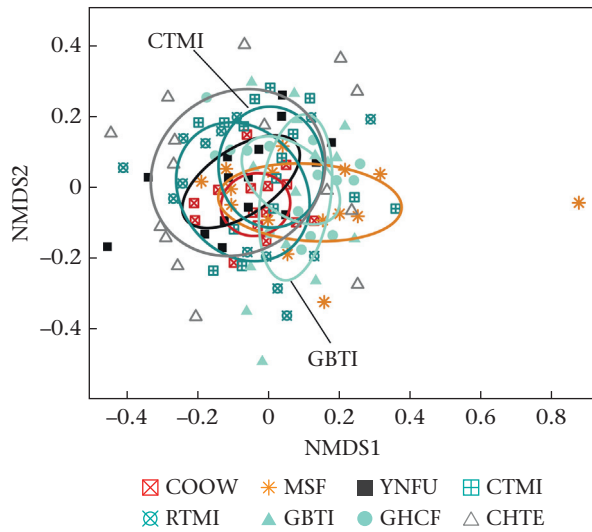


Figure 2. An NMDS visualization of the communities that responded to the different treatments. Colour represents type of treatment as in Fig. 1; symbols are different for all eight treatments. Each point represents the composition of the species that responded to two trials of that treatment at each playback location ($N = 15$ playback locations). The compositional data include abundances (numbers of individuals of each species). Ellipses represent 1 SD from the centroid of each treatment.

Comparison of Responding Species and Their Traits

We found 14/72 species showed nonrandom responses to the eight treatments. However, the result was affected by sample sizes, because species that responded rarely had less possibility to show a significant effect. For example, the species that had the fewest responses but still showed a significant nonrandom pattern responded six times. Yet 35 species responded less than six times and hence could not show a significant response. For species that responded at least 10 times, 12/22 were nonrandom. Of the 14 species with nonrandom response patterns, 11 responded most to the predator COOW treatment and six species responded least to nonflocking CHTE. Hence these patterns support the species-richness results given above (see Fig. 1).

Abundance was highly correlated with response (Fig. 3a; $t = 3.46$, $P = 0.001$, $R^2_{\text{adj}} = 0.17$). We therefore adjusted response by dividing it by the abundance estimate for each species. The abundance-adjusted response was not influenced by flocking propensity (Fig. 3b; $t = -0.15$, $P = 0.88$) but was influenced by the

following index (Fig. 3c; $t = 2.30$, $P = 0.029$, $R^2_{\text{adj}} = 0.12$). In the flocking propensity and following index analyses there was one outlier that was particularly strongly attracted although it had low abundance, yellow-bellied fairy-fantail, *Chelidorhynch hypoxanthus*. The responsiveness of this species makes sense, since it is a fly-catching species that follows MSFs, but even when this species was deleted from the analysis, the result for the following index was still similar ($t = 2.06$, $P = 0.049$, $R^2_{\text{adj}} = 0.10$). The body size of responding species and their diets were similar for all treatments (Figs A2 and A3).

DISCUSSION

Playback has been used by a number of studies to look at the roles of species in MSFs (Cordeiro et al., 2014; Goodale et al., 2012; Goodale & Kotagama, 2005b; Mönkkönen et al., 1996; Sullivan, 1984). It has also been used to probe species roles in some moving (birds that feed on army ants, Batcheller, 2017; Martínez et al., 2018) and nonmoving aggregations, such as frugivorous birds in fruiting trees (Gu et al., 2017), or in mobbing (Forsman & Mönkkönen, 2001; Hurd, 1996; Langham et al., 2006). A central question of interpretation is whether, when one simulates a species and heterospecifics are attracted, the heterospecifics came in order to associate with that species, or whether they were using that species as an indication of an MSF or aggregation (Goodale et al., 2020). In our experiment we included a treatment simulating the whole MSF system, and a range of species that participate inside it (a primary leader, two secondary leaders, two followers), as well as a stimulus, a predator call, that elicits mobbing. Here we describe three levels of results: (1) an MSF effect, in which MSF-follower species were attracted to most treatments that included regular flock participants (both leaders and followers), as if they were searching for an MSF rather than attracted to a particular leader; (2) some signs that responding birds were responding to differences in behaviour, and particularly in vigilance, between the simulated species; and (3) a general correspondence between mobbing and flocking in the species that were attracted to playback.

Aspects of the Methodology and Limitations

Before discussing our interpretation of the results of this experiment, we note a few special features of our methodology and some limitations. Most playback experiments focused on hetero-specific response have previously played all treatments at the same

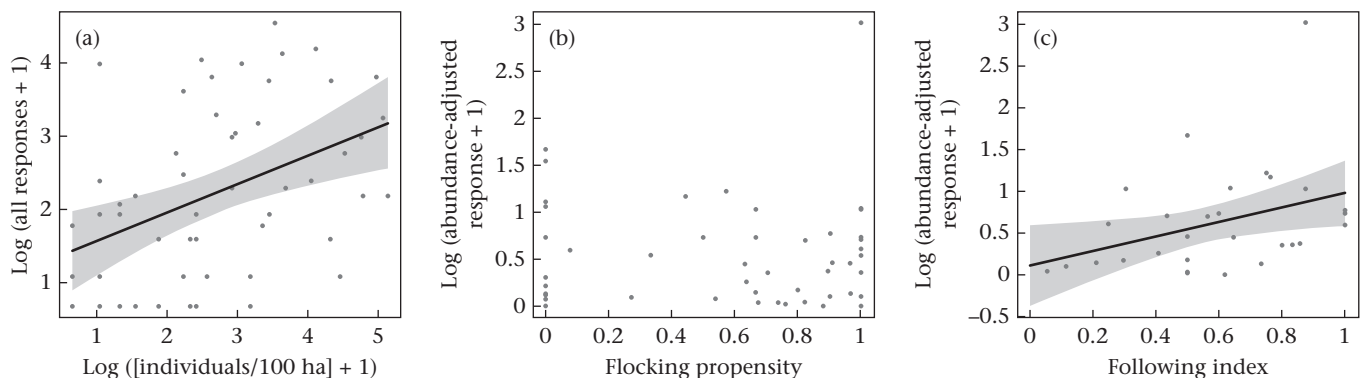


Figure 3. The influence of species' traits on their responses. Linear models relating: (a) each species' abundance (per 100 ha) to its total number of responses (in all treatments, $N = 54$ species), (b) species' flocking propensities (percentage of individuals observed in MSFs) to abundance-adjusted responses (total responses/bird per 100 ha; this analysis did not include trials of the predator COOW and nonflocking CHTE treatments; $N = 54$ species) and (c) species' following index (percentage of observations seen in the individual level leader/follower data in which the species was a follower; $N = 31$ species with three or more leader/follower observations) to abundance-related responses (again, predator COOW and nonflocking CHTE treatments excluded). The outlier with high responsiveness in (b) and (c) is the yellow-bellied fairy-fantail, *Chelidorhynch hypoxanthus*.

volume (e.g., [Martínez, Parra, et al., 2018](#), who used the number of calls and a stereo effect to differentiate treatments that simulated many individuals, compared to a few individuals, but played all treatments at the same amplitude). In some ways, such a strategy standardizes across treatments in that all trials might reach a similarly sized audience of birds a certain distance from the speaker. However, different frequencies of sounds travel different distances ([Brumm & Naguib, 2009](#); [Wiley & Richards, 1982](#)), and bird species differ in their natural amplitudes. Therefore, in this experiment we measured typical amplitudes of the simulated species and adjusted the playback of the different treatments so that differences in amplitude between the species were preserved.

We also included some factors in the analysis in novel ways. We investigated whether the number of conspecific responses (number of individuals) could have amplified the effects of the playback and found that it did not. We included the tape exemplar as a random factor in the GLMM modelling, thus incorporating the effect of playback tape in the statistical design ([Kroodsmma et al., 2001](#)). However, as it did not explain any variance, it was removed from the final models.

The experiment also had some limitations, however. We would have liked to include the seasonality of the experiment as another random factor (i.e. incorporate, for any trial, which round of sampling it was in), but the fact that we did all the predator COOW trials in the second year did not allow this. In our supplemental analyses, we investigated whether round or year influenced response when excluding the owl treatment and found that while these temporal factors did matter somewhat (the later rounds in the second year had less response, perhaps indicating some habituation), there was no interaction between the temporal factors and treatment. Hence, we can be confident that the relative differences between treatments did not change seasonally within or between years.

Comparison of the Number of Species that Responded

The pattern of response of heterospecifics suggests that they were attracted to an indication of an MSF and not to specific leader species, contrary to our hypotheses. As expected, response to most of the flocking species was higher than to the nonflocking control (with the largest exception being secondary leader RTMI, see below). Surprisingly, however, the playback treatment that simulated a whole MSF got no more response than those treatments that simulated different participating species separately, and leaders and followers were equally attractive. We selected the many simulated species to have dissimilar characteristics and suspected that different sets of respondents would be attracted to different simulated species in order to accrue specific benefits from them, with leader species being particularly attractive because of the many benefits they can offer. If this were true, the whole MSF treatment would have had the highest response, a sum of the other treatments. Such a 'sum effect' was seen in [Goodale and Kotagama \(2005b\)](#), in which a combination of the playback of two putative leader species attracted more species than the playback of either species separately. The lack of a sum effect in this study, combined with the similarities in response to the leaders and followers, suggests that species for the most part were attracted to all treatments because they sounded like an MSF. However, we must acknowledge that the tapes for the whole MSF treatment, which were selected because they had a high number of species calling simultaneously, did not necessarily include the same species that were separately simulated (5/5 had secondary leader RTMI, 2/5 had secondary leader CTMI and 1/5 had primary leader YNFU, with the other simulated species not included; [Table A2](#)), and this might have lessened response to that treatment (e.g. secondary leader RTMI was particularly not attractive to other species). It is also

possible that response to playback should not be expected to be additive, as a secondary leader might be attractive by itself but not attractive when the primary leader is there. Nevertheless, in our results birds came to both leaders and followers, and if they were coming to the playback because of specific benefits from these various species, we would still predict higher response for the whole flock treatment, which included many other follower species ([Table A2](#)). Another potential explanation of evenness in birds' responses to the different treatments would be that heterospecific attraction was just random. However, we show that this was not the case (primarily through the differential response to the control), as 12 of the 22 most frequent species showed nonrandom responses.

It was surprising that species' propensity for flocking was not an important factor predicting which species were attracted. To some extent this negative result might be because most of the species in this forest are usually in MSFs (ca. 60% of individuals in the forest were detected in MSFs), so there are few data from species with low propensities to join MSFs. It is also possible that some nonflocking species might synchronize their activity patterns with MSFs when the MSF is in their territory because it allows them to lower their predation risk temporarily, and to perform more risky behaviours (such as performing vocalizations associated with mate attraction, [Goodale et al., 2020](#)).

Leadership roles did influence which species were attracted to playback, with follower species being most responsive to the playback of MSF members. This result supports past research that followers are benefitting more than leaders in MSFs ([Dolby & Grubb, 1998](#); [Gentry et al., 2019](#); [Hino, 1998](#)). In particular, fly-catching species, such as the yellow-bellied fairy-fantail, which was the species with by far the highest abundance-adjusted response in this study, are following MSFs to consume insects the other birds disturb ([Satischandra et al., 2007](#); [Sridhar & Shanker, 2014](#)). Leader species may not particularly benefit from being in MSFs. Rather, the costs imposed on them by other species may not be large, so that it might not be worth trying to move away from the other species or drive them away ([Goodale et al., 2020](#)).

Indications that Some Characteristics of Simulated Species Mattered

We did see some indications in the results that behavioural traits of some of the simulated species made them especially attractive to heterospecifics. The highest correlations occurred among the predator COOW and primary leader YNFU treatments and between primary leader YNFU and follower GHCF treatments. Primary leader YNFU and follower GHCF are both species that have particularly high vigilance and provide information about predators. There are very many YNFU individuals per MSF, which gives them many eyes to detect a predator ([Goodale & Kotagama, 2005a](#)); it is difficult for human observers to approach this species without them alarm calling (for similar observations about closely-related species, see [Chen & Hsieh, 2002](#); [Jiang et al., 2020](#)). GHCF is a fly-catching (sallying) species, and it has been thought that such species have particularly high vigilance ([Jones & Sieving, 2019](#); [Martínez et al., 2016](#)). At high elevations in Sri Lanka, the usual sentinel species of MSFs, drongos, drop out, and GHCF, which is overall a very vocal species ([Shermila & Wickramasinghe, 2013](#)), is the species that makes the most alarm calls (E. Goodale, personal observation). So, as well as explaining the correlations between the responses to these treatments, these factors may explain why response to GHCF, a follower, was particularly high.

An unexpected species-specific finding was low response to secondary leader RTMI, which was a leader at the individual level, although not the flock level ([Table 1](#)). This species has a very particular foraging technique, running along tree branches like

nuthatches. Perhaps this specialization is the reason other species do not join it much; nuthatches, for example, are typically followers in MSFs (Dolby & Grubb, 1998). In contrast, leader species tend to have generalized feeding habits, so that their broad use of many vertical strata and their varied movement patterns allow many species to follow them (Hsieh & Chen, 2011).

Nevertheless, in discussing these traits of specific simulated species, we must acknowledge that playback responses to them were not significantly different compared to most other flocking members, even after a large sample ($N = 240$) of playback trials. As heterospecific attraction is in general weak, playback as a method may simply not have enough power to detect interactions between particular pairs of species, or even categories of species.

Comparison of Mobbing and Mixed-Species Flocking

We found a general similarity between mobbing elicited by the call of a small owl and attraction to species that participate in MSFs. The communities that responded to these very different stimuli were similar as shown by their overlap on the NMDS graph (Fig. 2). We must acknowledge here that the community that responded to the control nonflocking species was also similar in its mean NMDS scores. The mobbing and flocking communities (whole MSF treatment and three of the MSF members simulated), however, did have significantly less variance than the control treatment: the composition of the communities that responded to these treatments was similar at different locations, whereas what responded to the control was less predictable. Nevertheless, species that were attracted by the predator and those attracted by MSFs and participating species were all small birds (Fig. A2), with similar diets (Fig. A3). The correlation between species that responded to the predator COOW and whole MSF treatments was 0.6.

We had hypothesized some similarities between MSF and mobbing communities based on MSF leaders also being mobbing initiators, but the similarities between the communities were broader than we expected. The result emphasizes that both mobbing and flocking are adaptations that reduce predation risk for birds, and particularly for small birds. Mobbers tend to be those species that are vulnerable to the predator (and hence particularly small, Dutour et al., 2017b). Further, a recent meta-analysis showed that smaller species tend to be more often found in MSFs (Sainz-Borgo et al., 2018). One might expect the mobbing community to be different from the MSF one, if the species in MSFs are exclusive to MSFs and stay with them and are therefore not available on their territories to mob when a predator appears, but in our study site it seems that MSF members were found outside MSFs fairly often. This may occur because species that participate in MSFs may not be able to stay with the MSF when it goes outside their home range (Darrah & Smith, 2014) or birds may drop out of MSFs if the costs of moving at a nonoptimal speed are too high (Darrah & Smith, 2013; Hutto, 1988). If MSF-participating birds are outside an MSF and there is a predator present, they may then actively mob, and if they are fortunate, there may be enough other MSF-less individuals within hearing range to join them. Finally, note that mobbing and flocking are not ‘either-or’ phenomena: birds inside MSFs may also mob predators together (E. Goodale, personal observation).

Conclusions and Future Directions

In conclusion, heterospecifics came to most of the treatments that included MSF members, suggesting that they were attracted to an MSF. This result emphasizes that attraction to the simulation of leader or ‘keystone’ species could sometimes be due to those species being an indication of a group or a resource, and that experiments should try to test this hypothesis. Species that followed MSFs

were more attracted than leaders, reinforcing the idea that MSF followers benefit more from associating in MSFs than do leaders. Finally, there was a general correspondence between species that mob and species that join MSFs: the same species may participate in both these grouping phenomena in different situations. Future experiments comparing mobbing and flocking could use a variety of predator models to see whether displaying larger predators changes the similarity between the mobbing and flocking communities. This experiment was done at a site where flocking is dominant (ca. 60% of birds in MSFs); comparative experiments in communities where flocking is less strong might test whether flocking propensity could be a predictor of attraction in such environments.

Author Contributions

L.Z., R.-C.Q. and E.G. planned the study, L.Z., I.P. and K.L. carried out the fieldwork, L.Z. and E.G. analysed the data, L.Z. and E.G. wrote the first draft and all authors reviewed and approved the final manuscript.

Data Availability

The data are provided in the supplementary tables.

Conflict of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was funded by the National Natural Science Foundation of China, grant 31560119. We thank the Ailoshan Research Station of the Chinese Academy of Sciences for their permission to do the experiment and their help during the fieldwork. We are also grateful to Xiuwen Xu and Si Chen for help with the fieldwork, to Christos Mammides with advice about the statistics, and to two anonymous referees for their detailed and constructive comments that helped improve the manuscript.

Supplementary data

Supplementary material associated with this article can be found online at <https://doi.org/10.1016/j.anbehav.2021.08.012>.

References

- Banerjee, S., Schlaeppli, K., & van der Heijden, M. G. A. (2018). Keystone taxa as drivers of microbiome structure and functioning. *Nature Reviews Microbiology*, 16(9), 567–576. <https://doi.org/10.1038/s41579-018-0024-1>
- Batcheller, H. J. (2017). Interspecific information use by army-ant-following birds. *Auk*, 134(1), 247–255. <https://doi.org/10.1642/auk-16-93.1>
- Brumm, H., & Naguib, M. (2009). Environmental acoustics and the evolution of bird song. *Advances in the Study of Behavior*, 40, 1–33. [https://doi.org/10.1016/S0065-3454\(09\)40001-9](https://doi.org/10.1016/S0065-3454(09)40001-9)
- Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., & Thomas, L. (2001). *Introduction to distance sampling: estimating abundances of biological populations*. Oxford University Press.
- Carlson, N. V., Healy, S. D., & Templeton, C. N. (2017). A comparative study of how British tits encode predator threat in their mobbing calls. *Animal Behaviour*, 125, 77–92. <https://doi.org/10.1016/j.anbehav.2017.01.011>
- Carlson, N. V., Healy, S. D., & Templeton, C. N. (2020). What makes a ‘community informant’? Reliability and anti-predator signal eavesdropping across mixed-species flocks of tits. *Animal Behavior and Cognition*, 7, 214–246. <https://doi.org/10.26451/abc.07.02.13.2020>
- Chen, C.-C., & Hsieh, H. (2002). Composition and foraging behaviour of mixed-species flocks led by the Grey-cheeked Fulvetta in Fushan Experimental Forest, Taiwan. *Ibis*, 144, 317–330. <https://doi.org/10.1046/j.1474-919X.2002.00020.x>

- Chen, Q., Sun, J. Q., Song, L., Liu, W. Y., Yu, F. H., Li, S., Gong, H.-D., & Lu, H. Z. (2019). Trait acclimation of the clonal fern *Selliguea griffithiana* to forest epiphytic and terrestrial habitats. *Ecological Research*, 34(3), 406–414. <https://doi.org/10.1111/1440-1703.12002>
- Cordeiro, N. J., Borghesio, L., Joho, M., Monoski, T., & Mkongewa, V. (2014). Forest fragmentation in an African biodiversity hotspot impacts mixed-species bird flocks. *Biological Conservation*, 188, 61–71. <https://doi.org/10.1016/j.biocon.2014.09.050>
- Crowley, P. H., & Cox, J. J. (2011). Intraguild mutualism. *Trends in Ecology & Evolution*, 26(12), 627–633. <https://doi.org/10.1016/j.tree.2011.07.011>
- Curio, E. (1978). The adaptive significance of avian mobbing. I. Teleonomic hypotheses and predictions. *Zeitschrift für Tierpsychologie*, 48, 175–183. <https://doi.org/10.1111/j.1439-0310.1978.tb00254.x>
- Darrah, A. J., & Smith, K. G. (2013). Comparison of foraging behaviors and movement patterns of the Wedge-billed Woodcreeper (*Glyphorhynchus spirurus*) traveling alone and in mixed-species flocks in Amazonian Ecuador. *Auk*, 130, 629–636. <https://doi.org/10.1525/auk.2013.13088>
- Darrah, A. J., & Smith, K. G. (2014). Ecological and behavioral correlates of individual flocking propensity of a tropical songbird. *Behavioral Ecology*, 25(5), 1064–1072. <https://doi.org/10.1093/beheco/aru086>
- Diamond, J. (1987). Flocks of brown and black New Guinea birds: A bicoloured mixed-species foraging association. *Emu*, 87, 201–211. <https://doi.org/10.1071/MU9870201>
- Dolby, A. S., & Grubb, T. C., Jr. (1998). Benefits to satellite members in mixed-species foraging groups: An experimental analysis. *Animal Behaviour*, 56, 501–509. <https://doi.org/10.1006/anbe.1998.0808>
- Dugatkin, L. A., & Godin, J.-G. J. (1992). Prey approaching predators: A cost-benefit perspective. *Annales Zoologici Fennici*, 29, 233–252.
- Dutour, M., Lena, J.-P., & Lengagne, T. (2017a). Mobbing calls: A signal transcending species boundaries. *Animal Behaviour*, 131, 3–11. <https://doi.org/10.1016/j.anbehav.2017.07.004>
- Dutour, M., Lena, J. P., & Lengagne, T. (2017b). Mobbing behaviour in a passerine community increases with prevalence in predator diet. *Ibis*, 159(2), 324–330. <https://doi.org/10.1111/ibi.12461>
- Eguchi, K., Yamagishi, S., & Randrianasolo, V. (1993). The composition and foraging behaviour of mixed-species flocks of forest-living birds in Madagascar. *Ibis*, 135, 91–96. <https://doi.org/10.1111/j.1474-919X.1993.tb02814.x>
- Farine, D. R., Aplin, L. M., Sheldon, B. C., & Hoppitt, W. (2015). Interspecific social networks promote information transmission in wild songbirds. *Proceedings of the Royal Society B, Biological Sciences*, 282, 20142804. <https://doi.org/10.1098/rspb.2014.2804>
- Forsman, J. T., & Mönkkönen, M. (2001). Responses by breeding birds to hetero-specific song and mobbing call playbacks under varying predation risk. *Animal Behaviour*, 62, 1067–1073. <https://doi.org/10.1006/anbe.2001.1856>
- Gentry, K. E., Roche, D. P., Mugel, S. G., Lancaster, N. D., Sieving, K. E., Freeberg, T. M., & Lucas, J. R. (2019). Flocking propensity by satellites, but not core members of mixed-species flocks, increases when individuals experience energetic deficits in a poor-quality foraging habitat. *PLoS One*, 14(1), Article e0209680. <https://doi.org/10.1371/journal.pone.0209680>
- Goodale, E., & Beauchamp, G. (2010). The relationship between leadership and gregariousness in mixed-species bird flocks. *Journal of Avian Biology*, 41, 99–103. <https://doi.org/10.1111/j.1600-048X.2009.04828.x>
- Goodale, E., Beauchamp, G., Magrath, R., Nieh, J. C., & Ruxton, G. D. (2010). Inter-specific information transfer influences animal community structure. *Trends in Ecology & Evolution*, 25, 354–361. <https://doi.org/10.1016/j.tree.2010.01.002>
- Goodale, E., Beauchamp, G., & Ruxton, G. D. (2017). *Mixed-species animal groups: behavior, community structure and conservation*. Academic Press.
- Goodale, E., Goodale, U. M., & Mana, R. (2012). The role of toxic pitohuis in mixed-species flocks of lowland forest in Papua New Guinea. *Emu—Austral Ornithology*, 112, 9–16. <https://doi.org/10.1071/MU11026>
- Goodale, E., & Kotagama, S. W. (2005a). Alarm calling in Sri Lankan mixed-species bird flocks. *Auk*, 122, 108–120. <https://doi.org/10.1093/auk/122.1.108>
- Goodale, E., & Kotagama, S. W. (2005b). Testing the roles of species in mixed-species bird flocks of a Sri Lankan rainforest. *Journal of Tropical Ecology*, 21, 669–676. <https://doi.org/10.1017/S0266467405002609>
- Goodale, E., Nizam, B. Z., Robin, V. V., Sridhar, H., Trivedi, P., Kotagama, S. W., Padmalal, U. K. G. K., Perera, R., Pramod, P., & Vijayan, L. (2009). Regional variation in the composition and structure of mixed-species bird flocks in the Western Ghats and Sri Lanka. *Current Science*, 97, 648–663.
- Goodale, E., Sridhar, H., Sieving, K., Bangal, P., Colorado Zuluaga, G., Farine, D., Heymann, E. W., Jones, H. H., Krams, I., Martínez, A. E., Montano-Centellas, F., Muñoz, J., Srinivasan, U., Theo, A., & Shanker, K. (2020). Mixed company: A framework for understanding the composition and organization of mixed-species animal groups. *Biological Reviews*, 95, 889–910. <https://doi.org/10.1111/brv.12591>
- Greig-Smith, P. W. (1981). The role of alarm responses in the formation of mixed-species flocks of heathland birds. *Behavioral Ecology and Sociobiology*, 8, 7–10. <https://doi.org/10.1007/BF00302839>
- Gu, H., Chen, J., Ewing, H., Liu, X., Zhao, J., & Goodale, E. (2017). Heterospecific attraction to the vocalizations of birds in mass-fruiting trees. *Behavioral Ecology and Sociobiology*, 71(5), 82. <https://doi.org/10.1007/s00265-017-2312-6>
- Hetrick, S. A., & Sieving, K. E. (2012). Antipredator calls of tufted titmice and interspecific transfer of encoded threat information. *Behavioral Ecology*, 23(1), 83–92. <https://doi.org/10.1093/beheco/arr160>
- Hino, T. (1998). Mutualistic and commensal organization of avian mixed-species foraging flocks in a forest of western Madagascar. *Journal of Avian Biology*, 29, 17–24. <https://doi.org/10.2307/3677336>
- Hsieh, H., & Chen, C.-C. (2011). Does niche-overlap facilitate mixed-species flocking in birds? *Journal of Ornithology*, 152, 955–963. <https://doi.org/10.1007/s10336-011-0678-1>
- Hua, F., & Sieving, K. E. (2016). Understory avifauna exhibits altered mobbing behavior in tropical forest degraded by selective logging. *Oecologia*, 182(3), 743–754. <https://doi.org/10.1007/s00442-016-3695-1>
- Hurd, C. R. (1996). Interspecific attraction to the mobbing calls of black-capped chickadees (*Parus atricapillus*). *Behavioral Ecology and Sociobiology*, 38, 287–292. <https://doi.org/10.1007/s002650050244>
- Hutto, R. L. (1988). Foraging behaviour patterns suggest a possible cost associated with participation in mixed-species bird flocks. *Oikos*, 51, 79–83. <https://doi.org/10.2307/3565809>
- Hutto, R. L. (1994). The composition and social organization of mixed-species flocks in a tropical deciduous forest in western Mexico. *Condor*, 96, 105–118. <https://doi.org/10.2307/1369068>
- Jiang, D., Sieving, K. E., Meaux, E., & Goodale, E. (2020). Seasonal changes in mixed-species bird flocks and antipredator information. *Ecology and Evolution*, 10(12), 5368–5382. <https://doi.org/10.1002/ece3.6280>
- Jones, H., & Sieving, K. E. (2019). Foraging ecology drives social information reliance in an avian eavesdropping community. *Ecology and Evolution*, 9, 11584–11597. <https://doi.org/10.1002/ece3.5561>
- Julien, M., & Thiollay, J.-M. (1998). Multi-species territoriality and dynamic of neotropical forest understory bird flocks. *Journal of Animal Ecology*, 67, 227–252. <https://doi.org/10.1046/j.1365-2656.1998.00171.x>
- Kroodsma, D. E., Byers, B. E., Goodale, E., Johnson, S., & Liu, W.-C. (2001). Pseudoreplication in playback experiments, revisited a decade later. *Animal Behaviour*, 61, 1029–1033. <https://doi.org/10.1006/anbe.2000.1676>
- Langham, G. M., Contreras, T. A., & Sieving, K. E. (2006). Why pishing works: Titmouse (*Paridae*) scolds elicit a generalized response in bird communities. *Ecoscience*, 13(4), 485–496. [https://doi.org/10.2980/1195-6860\(2006\)13\[485:Wpwtips\]2.0.Co;2](https://doi.org/10.2980/1195-6860(2006)13[485:Wpwtips]2.0.Co;2)
- Latta, S. C., & Wunderle, J. M. (1996). The composition and foraging ecology of mixed-species flocks in pine forests of Hispaniola. *Condor*, 98, 595–607. <https://doi.org/10.2307/1369572>
- Martínez, A. E., & Gomez, J. P. (2013). Are mixed-species bird flocks stable through two decades? *American Naturalist*, 181, E53–E59. <https://doi.org/10.1086/685894>
- Martínez, A., Gomez, J. P., Ponciano, J. M., & Robinson, S. K. (2016). Functional traits, sociality and perceived predation risk in an Amazonian understory bird community. *American Naturalist*, 187, 607–619. <https://doi.org/10.1086/685894>
- Martínez, A. E., Parra, E., Muellerklein, O., & Vredenburg, V. T. (2018). Fear-based niche shifts in neotropical birds. *Ecology*, 99(6), 1338–1346. <https://doi.org/10.1002/ecy.2217>
- Martínez, A. E., Pollock, H. S., Kelley, J. P., & Tarwater, C. E. (2018). Social information cascades influence the formation of mixed-species flocks of ant-following birds in the Neotropics. *Animal Behaviour*, 135, 25–35. <https://doi.org/10.1016/j.anbehav.2017.10.024>
- Mönkkönen, M., Forsman, J. T., & Helle, P. (1996). Mixed-species foraging aggregations and heterospecific attraction in boreal bird communities. *Oikos*, 77, 127–136. <https://doi.org/10.2307/3545592>
- Moynihan, M. (1962). The organization and probable evolution of some mixed-species flocks of Neotropical birds. *Smithsonian Miscellaneous Collections*, 143, 1–140.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effect models. *Methods in Ecology and Evolution*, 4, 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>
- Paine, R. T. (1969). A note on trophic complexity and community stability. *American Naturalist*, 103, 91–93. <https://doi.org/10.1086/282586>
- Paine, R. T. (1995). A conversation on refining the concept of keystone species. *Conservation Biology*, 9, 962–964. <https://doi.org/10.1046/j.1523-1739.1995.09040962.x>
- Sainz-Borgo, C., Koffler, S., & Jaffe, K. (2018). On the adaptive characteristics of bird flocks: Small birds form mixed flocks. *Ornitologia Neotropical*, 29(1), 289–296.
- Satischandra, S. H. K., Kudavidanage, E. P., Kotagama, S. W., & Goodale, E. (2007). The benefits of joining mixed-species flocks for a sentinel nuclear species, the Greater Racket-tailed Drongo *Dicrurus paradiseus*. *Forktail*, 23, 145–148.
- Shermila, W. G. D. D. M., & Wickramasinghe, S. (2013). Composition of mix species foraging flocks of birds in Riverstan of montane region, Sri Lanka. *Journal of Tropical Forestry and Environment*, 3, 55–63. <https://doi.org/10.31357/jtfe.v3i1.1123>
- Sridhar, H., & Shanker, K. (2014). Using intra-flock association patterns to understand why birds participate in mixed-species foraging flocks in terrestrial habitats. *Behavioral Ecology and Sociobiology*, 68, 185–196. <https://doi.org/10.1007/s00265-013-1633-3>
- Sullivan, K. A. (1984). Information exploitation by downy woodpeckers in mixed-species flocks. *Behaviour*, 91, 294–311. <https://www.jstor.org/stable/4534393>
- Thiollay, J.-M. (2003). Comparative foraging behavior between solitary and flocking insectivores in a neotropical forest: Does vulnerability matter? *Ornitologia Neotropical*, 14, 47–65.
- Wiley, R. H., & Richards, D. G. (1982). Adaptations for acoustic communication in birds: Sound transmission and signal detection. In D. E. Kroodsma, & E. H. Miller (Eds.), *Acoustic communication in birds* (Vol. 2, pp. 131–181). Academic Press.

- Worm, B., & Paine, R. T. (2016). Humans as a hyperkeystone species. *Trends in Ecology & Evolution*, 31(8), 600–607. <https://doi.org/10.1016/j.tree.2016.05.008>
- Zhou, L., Peabotuwage, L., Gu, H., Jiang, D., Hu, G., Jiang, A., Mammides, C., Zhang, M., Quan, R.-C., & Goodale, E. (2019). The response of mixed-species flocks to anthropogenic disturbance and elevational change in southwest China. *Condor*, 121(3), duz028. <https://doi.org/10.1093/condor/duz028>
- Zhu, H. (2016). Discussion on the origin of mid montane wet evergreen broad-leaved forest in Yunnan. *Plant Science Journal*, 34, 715–723.
- Zou, F., Jones, H., Colorado Z, G. J., Jiang, D., Lee, T.-M., Martínez, C., Sieving, K., Zhang, M., Zhang, Q., & Goodale, E. (2018). The conservation implications of mixed-species flocking in terrestrial birds, a globally-distributed species interaction network. *Biological Conservation*, 224, 267–276. <https://doi.org/10.1016/j.biocon.2018.06.004>

Appendix

Methods

The objective of these analyses was to investigate the effects of seasonality and year on our results. Specifically, we wanted to know whether the results differed between the three rounds of playback conducted each year (early, middle or late), or between 2017 and 2018. Because all predator COOW trials were conducted in 2018, we could not incorporate these factors into the main analysis. Here we subdivided the data set in two different ways. First, we looked only at 2018 results (Model A1); for this analysis, the predator trials were divided into three categories based on when they were conducted (early, middle and late), and we used half the data set, sampling equally for each of the rounds. Second, we looked at the results of both years, excluding the predator treatment, and running two models, one with round as the explanatory factor (Model A2) and one with year as the explanatory factor (Model A3).

These GLMM models were conducted in a way similarly to those described in the main text. We included both the temporal factor (round or year) and the interaction between treatment and that temporal factor (treatment*round or treatment*year). Models were simplified as described in the main text.

We were also interested in seeing the effect of seasonality on the NMDS analysis, because that analysis showed the predator treatment in particular to have little variance, perhaps because all the trials were conducted in 1 year. So we ran an NMDS on data from 2018 only, again using half the predator trials and balancing them across seasons. In a first analysis using the responding community at each trial as a data point, the NMDS was too widely dispersed to visualize. Therefore, we did this analysis by considering a community to be all the individuals that responded to each playback tape (and hence $N = 5$ per treatment, except that nonflocking CHTE had a sample size of four, because there was no response to one exemplar).

Results

For the analysis from 2018 only, neither the interaction of treatment and round nor the factor of round was significant (Table A3). Multiple comparisons among treatments were qualitatively similar to those shown in the main text (Fig. A4).

For the results from both years excluding the predator treatment, again the interactions between treatment and the temporal factors (round or year) were not significant. However, round itself was significant (see Table A3); multiple comparisons among rounds showed that the second round tended to have higher response than the sixth round ($P = 0.06$; Fig. A5). Similarly, the first year of data had higher responses than the second year (see Table A3). Multiple comparisons among treatments in an analysis that included year

Table A1

The measurements of the amplitude (sound pressure level) of the vocalizations of the different species that were simulated

Species	Amplitude at 5 m (dB)
YNFU	46.8 ± 1.7
CTMI	45.5 ± 1.3
RTMI	54.6 ± 1.3
GBTI	51.0 ± 6.5
GHCF	57.5 ± 1.3
CHTE	54.2 ± 3.5

Measurements were made at 5 m, using a CEM DT-850 sound pressure level (SPL) meter, with A weighting and fast response settings. Sample sizes for each species = 5; mean values shown ± SD. YNFU = Yunnan fulvetta, *Alcippe fratercula*, CTMI = chestnut-tailed minla, *Actinodura strigula*, RTMI = red-tailed minla, *Minla ignotincta*, GHCF = grey-headed canary-flycatcher, *Culicicapa ceylonensis*, GBTI = green-backed tit, *Parus monticolus*, CHTE = chestnut-headed tesia, *Cettia castaneocoronata*.

Table A2

Species whose vocalizations were included in the five MSF exemplars (tapes A-E)

English name	Scientific name	Total no. of species
Tape A		7
Yunnan fulvetta	<i>Alcippe fratercula</i>	
Stripe-throated yuhina	<i>Yuhina gularis</i>	
Chestnut-tailed minla	<i>Actinodura strigula</i>	
Red-tailed minla	<i>Minla ignotincta</i>	
Mountain tailorbird	<i>Phyllergates cucullatus</i>	
Davison's leaf warbler?	<i>Phylloscopus intensior?</i>	
Yellow-bellied fairy-fantail	<i>Chelidorhynch hypoxanthus</i>	
Tape B		7
Whiskered yuhina	<i>Yuhina flavicollis</i>	
Crested finchbill	<i>Spizixos canifrons</i>	
Chestnut-tailed minla	<i>Actinodura strigula</i>	
Red-tailed minla	<i>Minla ignotincta</i>	
Bay woodpecker	<i>Blythipicus pyrrhotis</i>	
Davison's leaf warbler?	<i>Phylloscopus intensior?</i>	
White-tailed nuthatch	<i>Sitta himalayensis</i>	
Tape C		6
Red-tailed minla	<i>Minla ignotincta</i>	
Davison's leaf warbler	<i>Phylloscopus intensior</i>	
Gould's sunbird?	<i>Aethopyga gouldiae?</i>	
Yellow-cheeked tit?	<i>Macholophus spilnotus?</i>	
Yellow-bellied fairy-fantail	<i>Chelidorhynch hypoxanthus</i>	
Black-faced warbler	<i>Abroscopus schisticeps</i>	
Tape D		6
Red-tailed minla	<i>Minla ignotincta</i>	
Stripe-throated yuhina	<i>Yuhina gularis</i>	
Yellow-bellied fairy-fantail	<i>Chelidorhynch hypoxanthus</i>	
Yellow-cheeked tit	<i>Macholophus spilnotus</i>	
Fire-breasted flowerpecker?	<i>Dicaeum ignipectus?</i>	
?	?	
Tape E		6
Red-tailed minla	<i>Minla ignotincta</i>	
Davison's leaf warbler?	<i>Phylloscopus intensior?</i>	
Chestnut-crowned warbler	<i>Phylloscopus castaneiceps</i>	
Yellow-bellied fairy-fantail	<i>Chelidorhynch hypoxanthus</i>	
Rufous-winged fulvetta	<i>Schoeniparus castaneiceps</i>	
?	?	

Species that could not be identified are indicated by a question mark.

were qualitatively similar with those shown in the main text (e.g. Fig. A6 for the 2-year round analysis).

The NMDS on 2018 data was consistent with the one in the main analysis in that the variance for the predator COOW and the whole MSF treatments was particularly small (Fig. A7). It was different, however, in that the variation for nonflocking CHTE was not larger than for other treatments. One must remember, however, that this NMDS had to be run with a small data set, and nonflocking CHTE had only four data points.

Table A3

Results of models that incorporate temporal factors (round within year, different years)

Models	Fixed effect(s)	R^2m	R^2c	Factor	χ^2	df	P	Comparison	Z	P						
Model A1	Treatment + Treatment*Season + Season + Conspecific	0.44	0.49	Treatment	34.32	7	<0.001									
				Season	2.45	2	0.29									
				Conspecific	0.001	1	0.97									
				Treatment*Season	11.61	14	0.64									
	Treatment + Treatment*Season + Season	0.44	0.49	Treatment	34.38	7	<0.001									
				Season	2.45	2	0.29									
				Treatment*Season	11.70	14	0.63									
	Treatment + Season + Conspecific	0.37	0.43	Treatment	78.10	7	<0.001									
				Season	4.18	2	0.12									
				Conspecific	0.11	1	0.74									
	Treatment + Season	0.37	0.43	Treatment	78.72	7	<0.001									
				Season	4.15	2	0.13									
	Treatment	0.37	0.43	Treatment	78.89	7	<0.001	MSF–COOW	–3.24	0.03						
								YNFU–COOW	–4.28	<0.001						
								CTMI–COOW	–4.64	<0.001						
								RTMI–COOW	–6.01	<0.001						
								GBTI–COOW	–5.62	<0.001						
								GHCF–COOW	–3.95	0.00188						
								CHTE–COOW	–6.54	<0.001						
								YNFU–MSF	–1.03	0.97						
								CTMI–MSF	–1.52	0.79						
								RTMI–MSF	–3.20	0.03						
								GBTI–MSF	–2.58	0.16						
								GHCF–MSF	–0.66	1.00						
								CHTE–MSF	–4.02	0.001						
								CTMI–YNFU	–0.53	1.00						
								RTMI–YNFU	–2.28	0.30						
								GBTI–YNFU	–1.61	0.73						
								GHCF–YNFU	0.37	1.00						
								CHTE–YNFU	–3.19	0.03						
								RTMI–CTMI	–1.76	0.64						
								GBTI–CTMI	–1.08	0.96						
								GHCF–CTMI	0.89	0.99						
								CHTE–CTMI	–2.70	0.12						
								GBTI–RTMI	0.72	1.00						
								GHCF–RTMI	2.62	0.14						
								CHTE–RTMI	–1.04	0.97						
								GHCF–RTMI	1.97	0.49						
CHTE–RTMI								–1.74	0.65							
CHTE–RTMI								–3.50	0.01							
Model A2	Treatment + Round + Treatment*Round + Conspecific	0.32	0.36	Treatment	3.67	6	0.16									
				Round	1.35	5	0.93									
				Conspecific	0.61	1	0.43									
				Treatment*Round	34.31	30	0.27									
	Treatment + Round + Treatment*Round	0.31	0.36	Treatment	6.15	6	0.41									
				Round	1.44	5	0.92									
				Treatment*Round	34.80	30	0.25									
	Treatment + Round + Conspecific	0.20	0.25	Treatment	36.87	6	<0.001									
				Round	13.84	5	0.02									
				Conspecific	1.37	1	0.24									
	Treatment + Round	0.20	0.25	Treatment	37.08	6	<0.001	YNFU–MSF	–0.45	1						
								CTMI–MSF	0.05	1						
								RTMI–MSF	–2.94	0.05						
								GBTI–MSF	–2.56	0.14						
								GHCF–MSF	0.31	1.00						
								CHTE–MSF	–4.20	<0.001						
								CTMI–YNFU	0.50	1						
								RTMI–YNFU	–2.52	0.15						
								GBTI–YNFU	–2.11	0.34						
								GHCF–YNFU	0.76	0.99						
								CHTE–YNFU	–3.78	0.003						
								RTMI–CTMI	–2.99	0.04						
								GBTI–CTMI	–2.61	0.12						
								GHCF–CTMI	0.26	1						
								CHTE–CTMI	–4.25	<0.001						
								GBTI–RTMI	0.47	1						
								GHCF–RTMI	3.24	0.02						
								CHTE–RTMI	–1.28	0.86						
								GHCF–GBTI	2.88	0.06						
								CHTE–GBTI	–1.78	0.56						
								CHTE–GHCF	–4.50	<0.001						
								Round			14.19	5	0.01445	Second–First	0.04	1
														Third–First	–0.39	1.00
														Fourth–First	–1.79	0.46

Table A3 (continued)

Models	Fixed effect(s)	R^2m	R^2c	Factor	χ^2	df	P	Comparison	Z	P
Model A3	Treatment + Year + Year*Treatment + Conspecific	0.22	0.28	Treatment	17.49	6	0.008	Fifth–First	–1.53	0.64
				Year	0.09	1	0.77	Sixth–First	–2.40	0.15
				Conspecific	1.47	1	0.23	Third–Second	–0.46	1.00
				Treatment*Year	6.42	6	0.38	Fourth–Second	–1.45	0.69
								Fifth–Second	–1.84	0.43
								Sixth–Second	–2.74	0.07
								Fourth–Third	–1.04	0.90
								Fifth–Third	–1.25	0.81
								Sixth–Third	–2.20	0.23
								Fifth–Fourth	–0.08	1
	Treatment + Year + Treatment*Year	0.21	0.29	Treatment	17.73	6	0.007	Sixth–Fourth	–0.98	0.92
				Year	0.09	1	0.76	Sixth–Fifth	–0.99	0.92
				Treatment*Year	6.15	6	0.41			
	Treatment + Year + Conspecific	0.19	0.26	Treatment	36.93	6	<0.001			
				Year	12.29	1	<0.001			
				Conspecific	1.17	1	0.28			
	Treatment + Year	0.18	0.18	Year	12.81	1	<0.001	2018–2017	–3.58	<0.001
				Treatment	37.11	6	<0.001	YNFU–MSF	–0.45	1
								CTMI–MSF	0.07	1
								RTMI–MSF	–2.94	0.05
								GBTI–MSF	–2.55	0.14
								GHCF–MSF	0.32	1
								CHTE–MSF	–4.20	<0.001
								CTMI–YNFU	0.52	1
								RTMI–YNFU	–2.52	0.15
								GBTI–YNFU	–2.11	0.35
								GHCF–YNFU	0.77	0.99
								CHTE–YNFU	–3.78	0.003
								RTMI–CTMI	–3.01	0.04
								GBTI–CTMI	–2.62	0.12
								GHCF–CTMI	0.25	1
								CHTE–CTMI	–4.26	<0.001
								GBTI–RTMI	0.48	1
								GHCF–RTMI	3.25	0.02
								CHTE–RTMI	–1.27	0.86
								GHCF–GBTI	2.88	0.06
								CHTE–GBTI	–1.78	0.56
								CHTE–GHCF	–4.51	<0.001

For all three Poisson GLMM models, the number of respondent species per trial was the response variable. Model A1 used data of all treatments from 2018 only, and included the explanatory factor of round (early, middle and late); Model A2 used data from both years, but excluding the owlet treatment, and again included round as an explanatory factor (six rounds in total); Model A3 also used data from both years, but here the explanatory factor was year. All models also included interaction effects between treatment and the temporal factors before they were simplified.

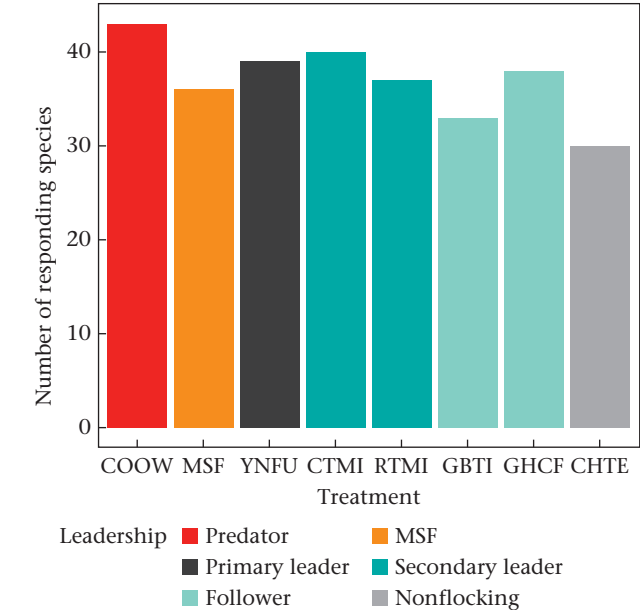


Figure A1. Total number of species that responded to the different treatments (30 trials per treatment). See Fig. 1 for treatment abbreviations.

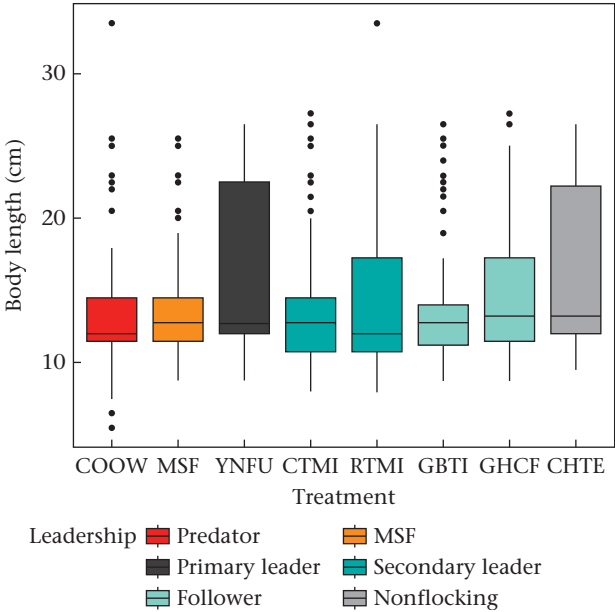


Figure A2. Body length (cm) of responding birds in the different treatments. For each treatment, we took the total list of responding individuals, and each individual was assigned the average body length of its species (data from Birds of the World, on-line edition, <https://birdsoftheworld.org/bow/home>). The box plot then shows the distribution of body lengths among responding individuals. See Fig. 1 for details of the box plot and treatment abbreviations.

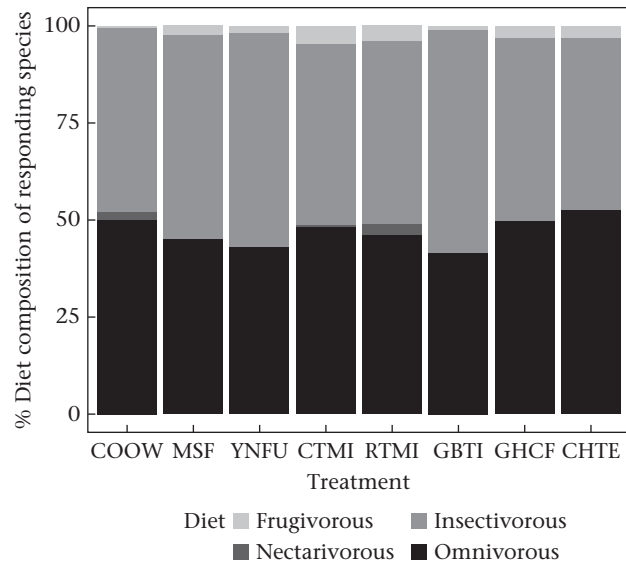


Figure A3. Diet of responding birds in the different treatments. For each treatment, we took the total list of responding individuals, and each individual was assigned one of four diets (frugivorous, insectivorous, nectarivorous, omnivorous; data from Birds of the World, on-line edition, <https://birdsoftheworld.org/bow/home>). The box plot then shows the distribution of diets among responding individuals. See Fig. 1 for details of the treatment abbreviations.

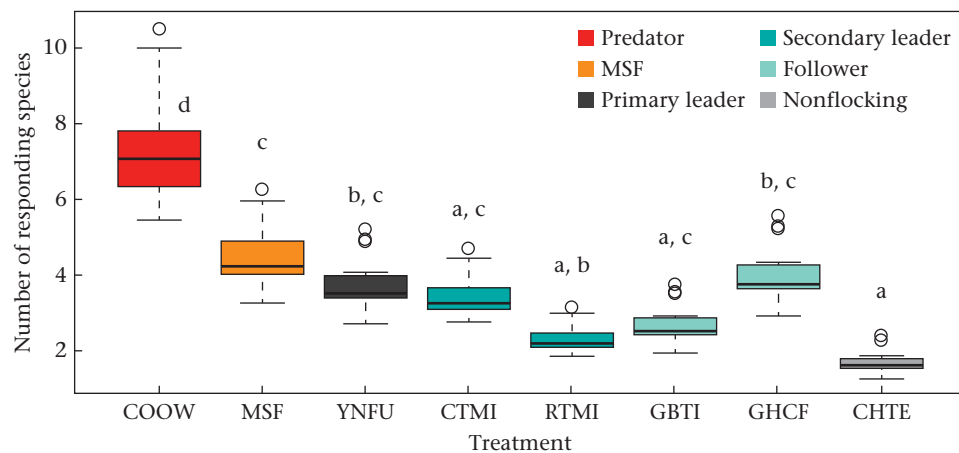


Figure A4. Number of species that responded per trial to the different treatments in 2018. The treatments are colour coded according to type (predator, MSF, primary leader, secondary leader, nonflocking). Treatments with the same letter were not significantly different from each other. See Fig. 1 for details of the box plot and treatment abbreviations.

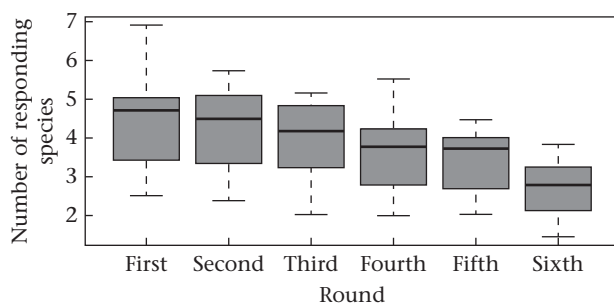


Figure A5. Number of species that responded per trial to the six rounds in both years (excluding the owl trials). Rounds 1–3 were in 2017 and rounds 4–5 were in 2018. See Fig. 1 for details of the box plot.

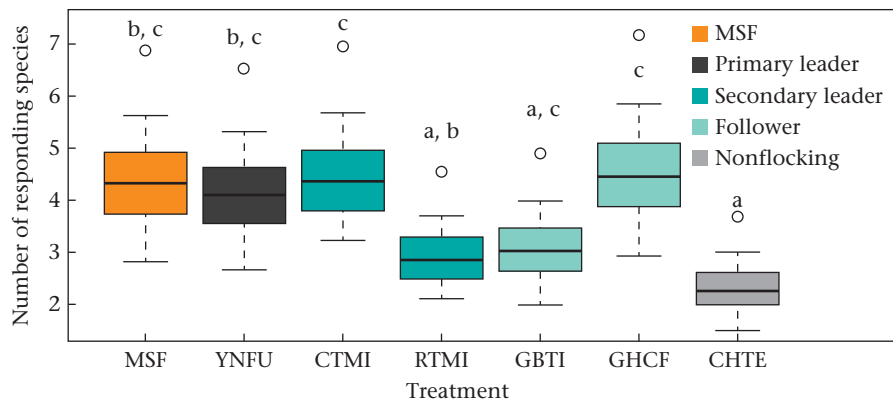


Figure A6. Number of species that responded per trial to the different treatments (without the owlet treatment) in an analysis that includes the factor year. The treatments are colour coded according to type (MSF, primary leader, secondary leader, follower, nonflocking). Treatments with the same letter were not significantly different from each other. See Fig. 1 for details of the box plot and treatment abbreviations.

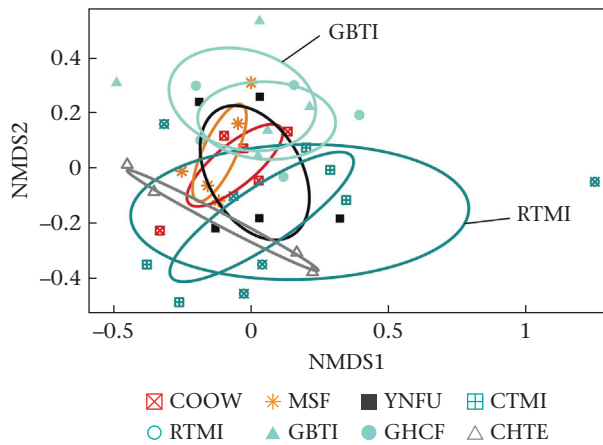


Figure A7. An NMDS visualization of the communities that responded to the different treatments in 2018. Colour represents type of treatment as in Fig. 1 in the main text; symbols are different for all eight treatments. Each point represents the composition of the species that responded to the same exemplar of a treatment ($N = 5$ exemplars). The compositional data included abundances (numbers of individuals of each species). Ellipses represent 1 SD from the centroid of each treatment.