Article:

Khwaja, N. orcid.org/0000-0003-1314-6984, Briskie, J.V. and Hatchwell, B.J. orcid.org/0000-0002-1039-4782 (2019) Individuality, kin similarity and experimental playback of contact calls in cooperatively breeding riflemen. New Zealand Journal of Zoology, 46 (4). pp. 334-347. ISSN 0301-4223

https://doi.org/10.1080/03014223.2019.1587477

This is an Accepted Manuscript of an article published by Taylor & Francis in New Zealand Journal of Zoology on 09/04/2019, available online: http://www.tandfonline.com/10.1080/03014223.2019.1587477."

Individuality, kin similarity and experimental playback of contact calls in cooperatively breeding riflemen

Nyil Khwaja[1, 2]*, James V. Briskie[2] and Ben J. Hatchwell[1]

[1]*Department of Animal and Plant Sciences, University of Sheffield, Sheffield, United Kingdom*

[2]*School of Biological Sciences, University of Canterbury, Christchurch, New Zealand*

*Correspondence: Broome Bird Observatory, PO Box 1313, Broome 6725, Western Australia. Email nyil.khwaja@cantab.net

Individuality, kin similarity and experimental playback of contact calls in cooperatively breeding riflemen

Riflemen/tītipounamu (*Acanthisitta chloris*) are kin-based cooperatively breeding birds, which appear able to recognise their relatives. Here, we investigate the potential for vocalisations to act as recognition cues in riflemen. We identified an appropriate contact call and recorded it at the nest from 19 adult riflemen. Measurements of call characteristics were individually repeatable. In addition, call similarity was significantly correlated with relatedness among all birds and among males. Thus, in principle, these contact calls contain sufficient information for individual recognition of familiar kin, and some assessment of relatedness between unfamiliar birds. To test whether riflemen responded differently to calls of kin, we broadcast calls of relatives and non-relatives as separate treatments in a playback experiment. Focal birds rarely responded aggressively or affiliatively, and their tendency to do so was unrelated to treatment. We conclude that zip calls are suitable kin recognition cues, but whether they are used as such remains unknown.

Keywords: *Acanthisitta chloris*, bioacoustics, cooperative breeding, kin discrimination, kin recognition, kin selection, rifleman, vocal communication

Introduction

Avian cooperative breeding systems are characterised by the apparent altruism of birds that help each other raise offspring (Cockburn et al. 2017). A key explanation for this altruism is that helpers gain indirect fitness by rearing their close relatives (Hamilton 1964), particularly when they have limited opportunities to reproduce themselves (e.g. Emlen 1982). To maximise indirect fitness helpers must choose reliably to help their kin. This may be achieved if relatives occur predictably in space, or if helpers can actively recognise these relatives (Cornwallis et al. 2009).

In most cooperative breeders, helpers delay dispersal from their parents' territory and help to raise their siblings (Emlen 1982, Hatchwell 2009, Riehl 2013). In these cases, helpers may benefit by simply rearing any offspring born on that territory (Wright et al. 1999, Dickinson and Hatchwell 2004). However, the spectrum of social organisation among cooperatively breeding species is much broader than this (Cockburn et al. 2017). For example, in the colonies of Australian *Manorina* miners and southern African sociable weavers (*Philetairus socius*), pairs live in large groups without breeding territories, and prospective helpers are surrounded by birds that differ in their degree of relatedness (Clarke and Fitz-Gerald 1994, Van Dijk et al. 2015). Likewise, in the more dispersed 'kin neighbourhoods' of western bluebirds (*Sialia mexicana*) and long-tailed tits (*Aegithalos caudatus*), offspring disperse to breed before subsequently moving between territories to help if their own attempts fail (Dickinson et al. 2016a, Hatchwell 2016). If helping is kin-selected, an ability to recognise kin is crucial for individuals to make optimal helping decisions in either of these social structures.

Riflemen/tītipounamu (*Acanthisitta chloris*) have a social system best described as a kin neighbourhood. Riflemen may breed twice in a season, and juveniles from a first brood sometimes remain with their parents to help provision the second (Sherley 1990a). However, the contribution they make is equivocal: they provision at low rates and are not associated with improved survival of recipient broods (Preston, Briskie, et al. 2013, 2016). More regular and effective help comes from adult birds that have previously dispersed away from their natal territory but 'commute' back to this territory, or another belonging to a close relative, to help. These birds may be unpaired, failed breeders or even breeding concurrently (Preston, Briskie, et al. 2013). Adult helpers are observed at around a quarter of nests. They are close genetic relatives of the broods they help, and these broods show enhanced recruitment rates, indicating that helping provides indirect fitness benefits (Preston, Briskie, et al. 2013, 2016). The movements that rifleman helpers make between territories and the consistency with which they help close relatives suggest they are able to recognise kin. This interpretation is supported by evidence that they actively avoid inbreeding (Preston 2012).

At least nine species of cooperatively breeding bird are thought to recognise their kin using vocalisations (McDonald and Wright 2011, Crane et al. 2015 and references within, Van Dijk et al. 2019). Vocal kinship cues may be individually distinct, and learned from association with nestmates and adults provisioning at the natal nest ('social relatives'). Alternatively, they may differ according to genetic differences between individuals. The latter have the advantages of permitting kin recognition between unfamiliar individuals, and of being robust to the influences of extra-pair paternity and conspecific brood parasitism, which reduce the average relatedness between social relatives. However, associatively learned cues appear more common (Riehl and Stern 2015, but see McDonald and Wright 2011).

Here, we investigate the hypothesis that the ability of riflemen to discriminate kin is based on the recognition of vocal cues. Vocal communication in riflemen has been the subject of three recent studies. Withers (2013) identified three elements, the 'zip', 'chuck' and 'pip', that make up the repertoire of rifleman contact calls. She showed that parameters from these calls differed significantly between separate subpopulations of riflemen on the North Island, but that adult birds did not respond differently to played-back calls from their own population as compared to those from other populations. Krull et al. (2009) demonstrated that adult contact calls and a nestling begging call contained ultrasonic components. They suggested that these play a role in communication or foraging; alternatively, they could exist as an adaptively neutral epiphenomenon. Preston (2012) also used a playback experiment, in this case testing whether breeding riflemen recognised the calls of their own juvenile offspring, or other adult kin in the population, using two pairwise experiments. She found some evidence that adults responded differently to the calls of their own juvenile offspring compared to unfamiliar juvenile calls. However, like Withers (2013), she found no difference in behavioural responses to different playback treatments of adult calls, leaving open the important question of how adult helpers recognise their kin.

We focus on the zip call for this study. We have observed that the pip is normally produced in a moderately distressed context (e.g. as a precursor to alarm calls), and chucks generally provide quiet and variable ‘punctuation’ between louder zips, so we considered the zip to be the most likely candidate call to be used to recognise kin in normal interactions. Zip calls are produced regularly in benign contexts such as pairs foraging together, and are not always interspersed with chucks. Furthermore, zip calls are uttered by adults before entering the nest (e.g. prior to replacing their partner during incubation), at a time when the signaller and receiver are unable to see each other, so they may encode information on identity. This behaviour continues into provisioning, meaning that nestlings could learn the zips of their parents in the nest. We recorded these calls in a standardised way, assessed their potential to encode information on identity and relatedness, and tested breeding birds’ responses to the calls of kin and non-kin using a playback experiment.

Material and methods

Fieldwork for this study was carried out on an individually colour-ringed, nestbox population of riflemen at Kowhai Bush, Kaikōura, between September 2014 and January 2015. The site is a temperate regenerating scrub forest dominated by kānuka (*Kunzea ericoides*). The study period corresponds to the rifleman breeding season. All riflemen in the population were genotyped at 16 polymorphic microsatellite loci (Preston, Dawson, et al. 2013, Khwaja et al. 2018).

Recording calls

We used a standardised procedure to record zip calls from breeding riflemen while they were incubating eggs. Eggs are incubated by both parents, but not helpers, who have not been observed visiting nests before eggs hatch (Sherley 1990b, Preston, Briskie, et al. 2013). We waited until neither parent was in the nest, and placed an Olympus ME15 tie-clip microphone inside the nestbox, clipped to the entrance of the nest chamber (riflemen build fully enclosed nests within nestboxes, rather than simply lining the box). The microphone was attached to an Olympus WS-812 recorder, which we left on top of the nestbox to record for about an hour. We started the recording simultaneously with a video camera mounted on a tripod, which filmed the nest entrance from c.10 m away. We watched the footage to determine when adults arrived at the nest to start an incubation bout. This often coincided with the utterance of a zip call, which we confirmed by checking a spectrogram of the audio recording generated using the program RavenPro 1.5 (Bioacoustics Research Program 2014). We isolated those calls that were unambiguously attributable to one individual, i.e. produced by an adult arriving when its partner was known not to be in the nest, and used them in the following analysis. In total, we recorded 156 zip calls from 19 adult riflemen.

Call individuality

Example spectrograms of rifleman zip calls are shown in Fig. 1. We used RavenPro’s selection tool to measure zip call parameters. For each call, we measured the peak frequency (frequency at maximum amplitude); first quartile (Q1), central and third quartile (Q3) frequencies (frequency at time point when 25%, 50% and 75% of the energy in the call has been produced, respectively); interquartile frequency range (difference between Q1 and Q3 frequencies); interquartile duration (the time before which 75% of the call's energy occurs minus the time before which 25% of the call's energy occurs), and average slope of the peak frequency contour (the mean difference between peak frequency measurements at successive time points separated by 1 ms). We did not use the measurements of minimum or maximum frequency, total frequency range, or total duration, because these measurements were hard to define: calls generally faded from their lowest frequency at the beginning and to their highest frequency at the end (Fig. 1), so it was difficult to standardise maximum and minimum frequencies. By focusing on quartile-based measurements we circumvented this issue, as the faded areas in question contained a negligible proportion of each call's total energy.

Some of these variables were expected to correlate with each other, particularly the frequency measurements, so we simplified further using a principal component analysis in R 3.4.4 (R Core Team 2018), after scaling and centring each variable. The first three principal components (PC1, PC2 and PC3) explained 92% of the variation in call measurements and were used in further analysis. PC1 score was mainly influenced by peak, central, Q1 and Q3 frequencies. PC2 score was mainly influenced by interquartile frequency range, duration and the slope of the peak frequency contour. PC3 score was mainly influenced by interquartile duration and the slope of the peak frequency contour (Table 1).

We checked for an effect of sex on the structure of calls by fitting three linear mixed-effects models (LMMs) using the nlme package (Pinheiro et al. 2010). PC1, PC2 and PC3 scores were the respective response variables. Sex was fitted as a fixed effect and individual identity as a random effect. There was no significant effect of sex on PC1 score (LMM: $F_{1, 17} = 1.59$, $P = 0.224$), a marginally non-significant effect of sex on PC2 score (LMM: $F_{1, 17} = 4.00$, $P = 0.062$), and no significant effect of sex on PC3 score (LMM: $F_{1, 17} = 0.01$, $P = 0.944$).

We assessed whether zip calls could encode information on individual identity using two approaches. First, we calculated LMM-based repeatability (*R*) for each principal component in the rptR package (Stoffel et al. 2017), using likelihood ratio tests to assess statistical significance. Significantly positive values of *R* indicate that there is individual consistency in the character of interest. Because of the lack of a significant sex effect on principal component scores we analysed both sexes together, though because this non-significance was marginal for PC2, we checked the validity of *R* by calculating an adjusted repeatability (R_{adj}), accounting for the possibly confounding effect of sex. Second, we used a discriminant function analysis in the MASS package (Venables and Ripley 2002) to calculate the proportion of calls that could be assigned to the correct individuals based on their PC1, PC2 and PC3 scores. To test whether this proportion was greater than we would expect by chance, we shuffled the observed principal component scores 1,000 times such that they were randomly assigned to individuals. We ran the analysis on each of these simulated datasets and derived a 95% confidence threshold from the distribution of these ‘expected’ proportions.

Call similarity and kinship

We compared the calls described above using a dynamic time warping algorithm implemented in the program Luscinia (Lachlan et al. 2010). This technique optimally aligns two signals to calculate a dissimilarity score based on specified parameters; we compared the durations and dynamic fundamental frequencies of calls to obtain this score. Luscinia then provides an overall dissimilarity score between two individuals based on the pairwise comparisons between their individual calls. We used these calls as measures of ‘acoustic distance’ between the individuals we studied.

We created acoustic distance matrices using these measures for all individuals, and also separately for males and females in light of the marginal effect of sex on call structure.

We calculated Queller-Goodnight genetic relatedness estimates between each dyad using the program SPAGeDi (Hardy and Vekemans 2002); using this estimator, negative values of relatedness correspond to dyads that are less related than the population average, while parent-offspring and full sibling dyads have an expected relatedness of 0.5 (Queller and Goodnight 1989). We subtracted relatedness from one as a measure of 'genetic distance' (Leclaire et al. 2012). We tested the relationship between acoustic distance and genetic distance using the vegan package (Oksanen et al. 2011), using 5,000 permutations to obtain a P-value for the correlation coefficient (r). Significantly positive values of r indicate that more closely related individuals have more similar calls, accounting for the non-independence of repeated comparisons using the same individuals.

Playback experiment

We used the zip calls recorded and analysed as above to create short playback tracks using the program Audacity 2.0.6 (Audacity Team 2014). These tracks consisted of a single representative zip call for an individual repeated every 10 seconds, reflecting natural calling rates (N. Khwaja pers. obs.) with generated silence in between.

For each of 14 individuals tested, we aimed to conduct two treatments: one playback of a first-order adult relative (mother, father, sibling or offspring), and one playback of an unrelated individual. We were unable to standardise the sex and (for the kin treatment) exact relationship of playback individuals, because of limited available kin within the population. For 4 of these individuals we were only able to complete one treatment (see Results). The treatment was known to the experimenter (i.e. the experiment was not blinded). We carried out each treatment on separate days for each individual (mean number of days between treatments for the same individual = 2.70, SD = 2.06, range = 1–7), with a maximum of one treatment per day at each nest. The order in which individuals were presented with the kin and non-kin treatments was alternated between experiments. Individuals were tested while they were provisioning nestlings, between day 13 and day 23 of a breeding attempt (where day 0 is the day of hatching). In each treatment, an experimenter erected a pop-up camouflage tent approximately 10 m from the nest, when neither adult was present. The experimenter started the playback track when the focal individual returned to view on its own; the understorey was open throughout the areas of forest in which rifleman nests were present, so there was little bias in visibility between nests. The playback track was played from a laptop using Audacity, connected to a pair of speakers placed outside the tent. The track was repeated until the individual left the area, or 10 minutes had elapsed. The following aspects of the individual's behaviour were noted during this time: whether it moved towards the speaker; its starting distance from, and closest approach to, the speaker; whether it called in response to the recording, and if so the type of call, and whether it changed its behaviour from one type to another (categorised as foraging, preening, calling, staying still, approaching the speaker or flying away). We tested whether birds' closest approach was significantly different between the two treatments using a paired, two-tailed *t*-test in R, expecting birds to be more likely to make aggressive approaches to the speaker in the non-kin treatment. We tested whether the binary variables of producing zip calls, producing pip calls, moving towards the speaker or changing behavioural type were significantly different between the two treatments using McNemar's test. We expected zip calls to be associated with affiliative behaviour and thus more likely to be uttered in the kin treatment, and the other three variables to be associated with aggression and more likely in the non-kin treatment.

In contrast to species where responses to playback are uniformly aggressive approaches to the speaker, varying only in severity or duration (e.g. Akçay et al. 2013), when carrying out our experiment we noticed that riflemen sometimes appeared to approach the speaker aggressively and sometimes affiliatively. At other times responses appeared neutral or absent. We were concerned therefore that measures such as tendency to approach the speaker and closest approach to it may not have been meaningful directional metrics of response. In light of this we also gave responses a subjective score as ‘aggressive’ (1, e.g. approaching speaker and mobbing, or producing pip calls), ‘neutral’ (2, for no response or a response that was not obviously aggressive or affiliative e.g. staying still, preening or continuing previous behaviour), and ‘affiliative’ (3, responding to calls with zips or chucks). We had insufficient sample size to use a multinomial response or McNemar’s test, so we used a paired Wilcoxon rank sum test to evaluate our prediction that individuals would score higher on this affiliation index when presented with playbacks of their kin.

Results

We recorded and measured 156 zip calls from 19 adult riflemen (11 males and 8 females, mean = 8.21 calls per bird, SD = 6.55, range = 1–24). The mean, minimum and maximum values for measurements of each call parameter are given in Table 1.

Individuality of calls

Repeated measures within individuals were more similar than expected by chance for both PC1 (LMM-based repeatability: $R = 0.40 \pm 0.11$ SE, $P < 0.001$) and PC2 (LMM-based repeatability: $R = 0.31 \pm 0.11$ SE, $P < 0.001$), though not for PC3 (LMM-based repeatability: $R = 0.03 \pm 0.04$ SE, $P = 0.198$). Using an adjusted repeatability approach to account for the marginal effect of sex on PC2 gave an equivalent result (LMM-based repeatability: $R_{\mathrm{adj}} = 0.26 \pm 0.04$ SE, $P < 0.001$). Most of the raw measurements these scores were based on also showed significant individual repeatability (see Table 1). This repeatability of call parameters implies the potential for zip calls to encode information on individual identity. Discriminant function analysis assigned 26% of all calls to the correct individual, significantly more than expected by chance (range of success percentages from 1,000 randomisations = 3–22, $P < 0.001$). A greater proportion of calls were assigned correctly for males (35/95) when analysed separately to females (20/61), but this difference was not statistically significant (proportion test: $\chi^2 = 0.12$, df = 1, P = 0.730).

Call similarity and kinship

Acoustic and genetic distances showed a marginally significant positive correlation in all dyads (Mantel test: $r = 0.15$, P = 0.044; Fig. 2a). This correlation was stronger in MM dyads (Mantel test: $r = 0.26$, P = 0.042; Fig. 2b) and absent in FF dyads (Mantel test: $r = -0.01$, P = 0.535; Fig. 2c).

Playback experiment

We conducted both treatments of our playback experiment on 10 individuals (3 males and 7 females). We carried out one treatment on a further 4 males (3 with a non-kin treatment only, because they had no known first-order kin in the population, and 1 with a kin treatment only, because we had insufficient time to carry out a non-kin treatment). We used only the first 10 individuals for the paired analyses. The responses we recorded were not significantly different between the two treatments, although there was a marginally nonsignificant trend for birds being more likely to change their behaviour in the kin treatment (Table 2). Including the individuals for which we could only carry out one treatment, and using a Fisher's exact test to analyse the proportions of responses in each category, made this trend slightly stronger but still nonsignificant (7/11 changed behaviour in the kin treatment, 3/13 in the non-kin treatment, Fisher's exact test: P = 0.061).

According to our qualitative categorisation, in the kin treatment 2 individuals responded affiliatively, 3 responded aggressively and 6 responded neutrally. In the non-kin treatment, 1 individual responded aggressively and 12 responded neutrally. It is noteworthy that the most aggressive response we observed, i.e. mobbing of the speaker, was in response to the non-kin treatment. Nevertheless there was no evidence of a general preference for kin, with birds not responding more affiliatively towards them overall (paired Wilcoxon test: V = 2, n = 11 paired responses, P = 0.807; Fig. 3), and most responding neutrally to the experiment.

Discussion

We tested the potential role of the rifleman zip call in kin recognition. This is a short, simple contact call that appears to be readily used by birds when communicating with one another. By measuring the components of zip calls, we confirmed that they contained characteristics that were individually repeatable and enabled automated recognition at a rate that was significantly better than random. We also demonstrated that relatedness and call similarity were positively correlated overall and between males, but not between females. We used a playback experiment to test whether provisioning riflemen responded differently to the calls of their kin and non-kin. Although the most aggressive response was observed in a non-kin treatment, we found no general tendency for riflemen to behave more affiliatively when exposed to kin.

For a cue to be used to recognise individuals, repeated measures from the same individual should be more similar than expected by chance (Falls 1982). We confirmed this was the case for zip calls using repeatability tests and discriminant function analysis on measured parameters. Individual-specific call parameters have been commonly reported in birds, including in other cooperative breeders (Crane et al. 2015). However, they are not universal: for example, in contrast to ‘churr’ calls, the ‘tut’ calls of long-tailed tits are considered unlikely to be appropriate recognition cues because of a lack of between-individual variation (Sharp and Hatchwell 2005). The results of our analysis supported our inference from observing rifleman behaviour that zip calls were appropriate candidate cues to use to investigate recognition.

Although discriminant function analysis assigned significantly more calls to the correct individuals than expected by chance, this was still only successful in 26% of cases. This is considerably less than in many other studies, for example c.50% reported for long-tailed tit churr calls (Sharp and Hatchwell 2005) and 100% from a study of short-toed treecreeper (*Certhia brachydactyla*) calls (Bauer and Nagl 1992). The measurements we were able to take would not have captured all of the variation in rifleman zips, particularly given that we omitted some because of concerns over their accuracy (see Methods). The success rate is perhaps unsurprising in light of a study of riflemen on the North Island, in which discriminant function analysis assigned only 47% of zip calls to the correct subpopulation (Withers 2013); considering the isolation of these subpopulations, this level of differentiation is understandably greater than that found between individuals. The relative simplicity of zip calls is likely to make them more difficult to classify by acoustic analysis. The regularity with which riflemen use the zip call in social interactions may reduce recognition errors associated with this apparently imperfect cue, as they would usually have many samples of the call on which to base recognition. Riflemen do occasionally help non-relatives (c.10% of helping events), and we have not documented any fitness benefits of doing so (Khwaja 2017). It is possible that these are ‘mistakes’ driven by misclassifications of cues like the zip call.

Individually distinctive calls of social relatives can be learned through association and thus enable kin recognition between familiar birds (Sharp et al. 2005). In riflemen, such a cue would be reliable: an individual’s offspring, nestmates, social parents and parents’ future offspring are all genuine genetic relatives because of an absence of extra-pair paternity and intraspecific brood parasitism (Preston, Briskie, et al. 2013). Indeed, associatively learned kinship cues appear to be the norm in cooperatively breeding birds (Riehl and Stern 2015). However, purely learned cues do not permit recognition of unfamiliar kin. For this to be accomplished there must be more information content in the signal, for example, more genetically similar individuals having more similar call structures. We showed that this was the case in riflemen for all birds and among males, but not among females. Although there is considerable variation around the trend, assessment of relatedness by self-referent phenotype matching could therefore provide information about the likely kinship of unfamiliar birds. This would be potentially useful in making adaptive helping decisions; although we have not observed helping between birds that we knew from direct observation to be genetic but not social relatives, pedigree reconstruction suggests this is likely to have been the case for a minority of helpers (Khwaja 2017). Further, it would help birds avoid inbreeding when choosing a mate (Preston 2012). Recent analysis in western bluebirds suggests inbreeding can be effectively avoided through social recognition of familiar kin, in combination with female-biased dispersal (Dickinson et al. 2016b). However, genetic recognition may be more advantageous in riflemen, where dispersal of both sexes is limited (Preston 2012).

Although overall more genetically similar individuals had more similar zip calls, this correlation was stronger in males and absent between females. There is no obvious biological mechanism to produce this curious result, and it could be an artefact of sampling in an already small dataset. Nevertheless, it is noteworthy that it echoes findings from studies of both bell miners (*Manorina melanophrys*) and sociable weavers. In bell miners, vocal similarity closely matches genetic relatedness between helpers and male, but not female, breeders (McDonald and Wright 2011). In sociable weavers, the mean fundamental frequency of contact calls is more similar between more closely related males, but not females (Van Dijk et al. 2019). For both species, the pattern fits neatly into an adaptive explanation of the cooperative breeding system: both live in colonies where males show limited dispersal, and females are generally immigrants. Helpers (usually male) thus vary in their relatedness to breeding males, but are unrelated to most or all breeding females except their mothers. This makes call similarity a useful indicator of the indirect fitness gains available from helping a particular male. By contrast, in this rifleman population, although helping is also male-biased, males and females disperse similar distances and equivalent fitness gains are available to females helping their sisters and mothers, as males helping their brothers and fathers (Preston 2012), so there is no obvious adaptive basis for a male-specific effect. If this trend for genetic and vocal similarity to correlate only in males continues in studies of other birds, it may require explanation at a more general level than details of specific social structures.

The above results indicate that rifleman zip calls contain sufficient information to be useful as a cue for birds to recognise familiar and/or unfamiliar kin. Despite this, our playback did not support the hypothesis that riflemen responded differently to calls of kin and non-kin. One possible explanation is the low statistical power associated with the small sample of birds we were able to test: there was a trend for birds to change their behaviour more in the kin treatment, which may reflect a greater tendency to engage with their relatives. Another issue is that, due to limited available first-order kin in the population, we were unable to standardise the sex or relationship type of playback individuals. It is also possible that the experiment did not reflect accurately enough the context in which birds make helping decisions. Previous studies of other cooperative bird species have found that speakers were approached more closely, or more often, when broadcasting non-kin calls (e.g. Payne et al. 1998, Hatchwell et al. 2001, Keen et al. 2013), suggesting that sustained approach is a symptom of aggressive behaviour. This interpretation was probably not appropriate in our experiment. Riflemen moved towards the speaker in 13 of the 25 trials (7/11 kin treatments, 6/13 non-kin treatments), but these approaches ranged from the affiliative, in which the focal bird responded to broadcast zips with zips of their own, through seemingly neutral responses, to an aggressive approach which culminated in the focal bird mobbing the speaker. Because of this variation, tendency to and proximity of approach may not have been informative measures of response. For this reason, we also reported a qualitative assessment of focal birds' responses. However, they were often unresponsive to the experiment, and their behaviour was difficult to interpret: of 25 experimental trials, only 6 responses were clearly either affiliative or aggressive, and the type of response was not associated with the experimental treatment.

Overall, it appeared that playback as a simulated territorial intrusion did not generally provoke alarm from riflemen. Riflemen are mostly considered weakly territorial; we have witnessed disputes at territory boundaries and a tendency to remain within the confines of a territory, but aggression between conspecifics was rarely observed over the course of our study (N. Khwaja pers. obs.). Similar observations have been reported previously in our study population (Hunt and McLean 1993, Sherley 1994, Preston, Briskie, et al. 2013). Crane et al.'s (2015) experiments on another weakly territorial species, the chestnut-crowned babbler (*Pomatostomus ruficeps*), were also notable for their lack of aggressive responses by focal birds (but see Sharp et al. 2005). Crane et al. (2015) demonstrated a differential response by simultaneously broadcasting calls from group members and outsiders. In this scenario, the focal group moved affiliatively towards the call of their group member. Using a similar protocol in riflemen may be a more appropriate test for any future studies, given that responses to a single playback did not differ according to kinship status.

This study adds to two previous playback experiments using adult rifleman calls that have given negative results (Preston 2012, Withers 2013). This may have arisen from limitations in study design, or perhaps because the low responsiveness of riflemen has made successful experiments challenging to implement. Other possibilities are that vocalisations alone are insufficient stimuli for riflemen to behave naturally towards conspecifics, or that they are unable to recognise the calls of their kin. As discussed above, although zip calls were individually identifiable more than expected by chance, we classified them successfully in only 26% of cases. Similarly, although statistically significant, there is substantial variation around the trend for closer kin to have more similar calls, along with an apparent lack of such a relationship between females (Fig. 2). Thus, calls may be unreliable cues of kinship on their own, and receivers may have to incorporate information from other sources, such as olfactory or visual signals. The potential for riflemen (or to our knowledge any cooperatively breeding birds) to recognise one another visually has not been tested. Most birds have strong visual systems, but experimental study in this area has been neglected owing to the practical difficulties of identifying and systematically manipulating subtle individual visual characters (Nakagawa and Waas 2004, although see Whitfield 1986 for a study of individual visual recognition in a territorial wader). Addressing this gap represents a challenge for the future of avian kin recognition research. Another is to design experiments that replicate the key decisions for which birds need to (and from which we infer they can) recognise their kin. In riflemen these are the decision to help and the selection of a mate, rather than the tolerance of breeding birds to manipulations. The difficulties of this approach are clear, but it may bring us closer to understanding the mechanisms that drive and constrain kin-based cooperative breeding.

Acknowledgements: Thanks to Sarah Withers for helpful discussion of the study design, and Archie MacFarlane for discussion of the manuscript. Steph Preston contributed genotype data for four individuals in the genetic analysis. Genotyping was conducted at the Natural Environment Research Council (NERC) Biomolecular Analysis Facility at Sheffield (NBAF361), and we further thank Deborah Dawson, Gavin Horsburgh and Terry Burke for their advice and support. Two anonymous referees provided comments that improved the analysis and content of the manuscript. This research was funded by NERC (grant number 1233848) and the University of Sheffield. Fieldwork was carried out under Department of Conservation permit NM-34956-FAU, and approved by the University of Canterbury's Animal Ethics Committee.

References:

Audacity Team 2014. Audacity: free recording and editing software. https://audacityteam.org

Akçay C, Swift, RJ, Reed VA, Dickinson JL. 2013. Vocal kin recognition in kin neighborhoods of western bluebirds. Behav Ecol. 24: 898–905.

Bauer HG, Nagel W. 1992. Individual distinctiveness and possible function of song parts of short-toed treecreepers (*Certhia brachydactyla*) – evidence from multivariate song analysis. Ethology 91: 108–121.

Bioacoustics Research Program. 2014. Raven pro: interactive sound analysis software (version 1.5). Ithaca (NY): Cornell Lab of Ornithology.

Clarke MF, Fitz-Gerald GF. 1994. Spatial organisation of the cooperatively breeding bell miner *Manorina melanophrys*. Emu 94: 96–105.

Cockburn A, Hatchwell BJ, Koenig WD. 2017. Sociality in birds. In: Rubenstein DR, Abbot P, editors. Comparative social evolution. Cambridge (UK): Cambridge University Press; p. 320–353.

Cornwallis CK, West SA, Griffin AS. 2009. Routes to indirect fitness in cooperatively breeding vertebrates: kin discrimination and limited dispersal. J Evol Biol 22: 2245–2457.

Crane JMS, Pick JL, Tribe AJ, Vincze E, Hatchwell BJ, Russell AF. 2015. Chestnut-crowned babblers show affinity for calls of removed group members: a dual playback without expectancy violation. Anim Behav. 104: 51–57.

Dickinson JL, Hatchwell BJ. 2004. Fitness consequences of helping. In: Koenig WD, Dickinson JL, editors. Ecology and evolution of cooperative breeding in birds. Cambridge (UK): Cambridge University Press; p. 48–56.

Dickinson JL, Akçay C, Ferree ED, Stern CA. 2016a. Western bluebirds: lessons from a marginal cooperative breeder. In: Koenig WD, Dickinson JL, editors. Cooperative breeding in vertebrates: studies of ecology, evolution, and behavior. Cambridge (UK): Cambridge University Press, p. 19–38.

Dickinson JL, Akçay C, Ferree ED, Stern CA. 2016b. A hierarchical analysis of incest avoidance in a cooperative breeder. Behav Ecol. 27: 1132–1140.

Emlen ST. 1982. The evolution of helping. I. An ecological constraints model. Am Nat. 119: 29–39.

Hamilton WD. 1964. The genetical evolution of social behaviour. J Theor Biol. 7: 1–52.

Hardy OJ, Vekemans X. 2002. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes. 2: 618–620.

Hatchwell BJ. 2009. The evolution of cooperative breeding in birds: kinship, dispersal and life history. Phil Trans Roy Soc B. 364: 3217–3227.

Hatchwell BJ. 2016. Long-tailed tits: ecological causes and fitness consequences of redirected helping. In: Koenig WD, Dickinson JL, editors. Cooperative breeding in vertebrates: studies of ecology, evolution, and behavior. Cambridge (UK): Cambridge University Press, p. 39–57.

Hatchwell BJ, Ross DJ, Fowlie MK, McGowan A. 2001. Kin discrimination in cooperatively breeding long-tailed tits. Proc Roy Soc B. 268: 885–890.

Hunt GR, McLean IG. 1993. The ecomorphology of sexual dimorphism in the New Zealand rifleman *Acanthisitta chloris*. Emu 93: 71–78.

Keen SC, Meliza CD, Rubenstein DR. 2013. Flight calls signal group and individual identity but not kinship in a cooperatively breeding bird. Behav Ecol. 24: 1279–1285.

Khwaja N. 2017. Investment patterns and kinship cues in a cooperatively breeding bird [dissertation]. Sheffield (UK): University of Sheffield.

Khwaja N, Preston SAJ, Briskie JV, Hatchwell BJ. 2018. Testing the predictions of sex allocation hypotheses in dimorphic, cooperatively breeding riflemen. Ecol Evol. 7: 3693–3701.

Krull CR, Parsons S, Hauber ME. 2009. The presence of ultrasonic harmonics in the calls of the rifleman (*Acanthisitta chloris*). Notornis 56: 158–161.

Lachlan RF, Verhagen L, Peters S, Ten Cate C. 2010. Are there species-universal categories in bird song phonology and syntax? A comparative study of chaffinches (*Fringilla coelebs*), zebra finches (*Taenopygia guttata*), and swamp sparrows (*Melospiza georgiana*). J Comp Psychol. 124: 92–108.

Leclaire S, Merkling T, Raynaud C, Mulard H, Bessière J-M, Lhuillier E, Hatch SA, Danchin E. 2012. Semiochemical compounds of preen secretion reflect genetic make-up in a seabird species. Proc Roy Soc B. 279: 1185–1193.

McDonald PG, Wright J. 2011. Bell miner provisioning calls are more similar among relatives and are used by helpers at the nest to bias their effort towards kin. Proc Roy Soc B. 278: 3403–3411.

Nakagawa S, Waas J. 2004. ‘O sibling, where art thou?’ – a review of avian sibling recognition with respect to the mammalian literature. Biol Revs. 79: 101–119.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, et al. 2011. vegan: community ecology package. https://CRAN.R-project.org/package=vegan

Payne RB, Payne LL, Rowley I. 1988. Kin and social relationships in splendid fairy-wrens: recognition by song in a cooperative bird. Anim Behav. 36: 1341–1351.

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2010. nlme: linear and nonlinear mixed effects models. https://cran.r-project.org/web/packages/nlme/index.html

Preston SAJ. 2012. Routes to cooperation in the rifleman *Acanthisitta chloris* [dissertation]. Sheffield (UK): University of Sheffield.

Preston SAJ, Briskie JV, Burke T, Hatchwell BJ. 2013. Genetic analysis reveals diverse kin-directed routes to helping in the rifleman *Acanthisitta chloris*. Mol Ecol. 22: 5027-5039.

Preston SAJ, Dawson DA, Horsburgh GJ, Hatchwell BJ. 2013. Characterisation of microsatellite loci in the rifleman (*Acanthisitta chloris*, Acanthisittidae, Aves) and their predicted genome locations. Cons Gen Res. 5: 555–560.

Preston SAJ, Briskie JV, Hatchwell, BJ. 2016. Adult helpers increase the recruitment of closely related offspring in the cooperatively breeding rifleman. Behav Ecol. 16: 41–47.

R Core Team 2018. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Riehl C. 2013. Evolutionary routes to non-kin cooperative breeding in birds. Proc Roy Soc B. 280: 20132245.

Riehl C, Stern CA. 2015. How cooperatively breeding birds identify relatives and avoid incest: new insights into dispersal and kin recognition. BioEssays 37: 1303–1308.

Sharp SP, Hatchwell BJ. 2005. Individuality in the contact calls of cooperatively breeding long-tailed tits (*Aegithalos caudatus*). Behaviour 142: 1559–1575.

Sherley GH. 1990a. Co-operative breeding in riflemen (*Acanthisitta chloris*) benefits to parents, offspring and helpers. Behaviour 112: 1–22.

Sherley GH. 1990b. Relative parental effort during incubation in riflemen (*Acanthisitta chloris*). NZ J Zool. 20: 211–217.

Sherley GH. 1994. Co-operative parental care; contribution of the male rifleman (*Acanthisitta chloris*) to the breeding effort. Notornis 41: 71–81.

Stoffell MA, Nakagawa S, Schielzeth H. 2017. rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. Meth Ecol Evol. 8: 1639–1644.

Van Dijk RE, Covas R, Doutrelant C, Spottiswoode CN, Hatchwell BJ. 2015. Fine-scale genetic structure reflects sex-specific dispersal strategies in a population of sociable weavers (*Philetairus socius*). Mol Ecol. 24: 4296–4311.

Van Dijk RE, Khwaja N, Goldberg R, Argüelles Ticó A, Covas R, Doutrelant C, Hatchwell BJ. 2019. Sociable weavers use individually distinct vocalisations to recognise social group membership. In review.

Venables WN, Ripley BD. 2002. Modern applied statistics with S. New York (NY): Springer.

Withers S. 2013. Population variation of the North Island rifleman (*Acanthisitta chloris granti*): implications for conservation management [dissertation]. Auckland: University of Auckland.

Whitfield DP. 1986. Plumage variability and territoriality in breeding turnstone *Arenaria interpres*: status signalling or individual recognition? Anim Behav. 34: 1471–1482.

Wright J, Parker PG, Lundy KJ. 1999. Relatedness and chick feeding effort in the cooperatively breeding Arabian babbler. Anim Behav. 58: 779–785.

Table 1. Parameters measured from 156 rifleman zip calls. The LMM-based repeatability of calls from the same individual is shown for each parameter, along with its statistical significance in brackets (*** < 0.001, ** < 0.010, * < 0.050, . < 0.010, NS > 0.010). Contributions to each of the three major principal components (PC1, PC2 and PC3) are provided as percentages. These components respectively accounted for 55%, 22% and 14% of the total variation in the measured parameters.

Parameter	**Mean ± SE**	**Min.**	**Max.**	**Repeatability**	**% loading**		
					PC1	**PC2**	**PC3**

Q1 frequency (kHz)	8.84 ± 0.08	6.03	11.03	0.40 (***)	22	3	3
Central frequency (kHz)	9.24 ± 0.07	6.72	11.20	0.40 (***)	22	5	4
Q3 frequency (kHz)	9.67 ± 0.07	7.02	11.53	0.43 (***)	21	13	5
Peak frequency (kHz)	9.18 ± 0.08	5.86	11.20	0.35 (***)	22	2	2
IQ frequency range (kHz)	0.83 ± 0.03	0.17	2.24	0.29 (***)	5	35	4
IQ duration (ms)	7.20 ± 0.24	2.90	14.51	0.09 (NS)	8	20	40
Peak frequency contour mean slope (kHz/ms)	93.77 ± 4.27	-74.2	311.6	0.08 (.)	1	24	42

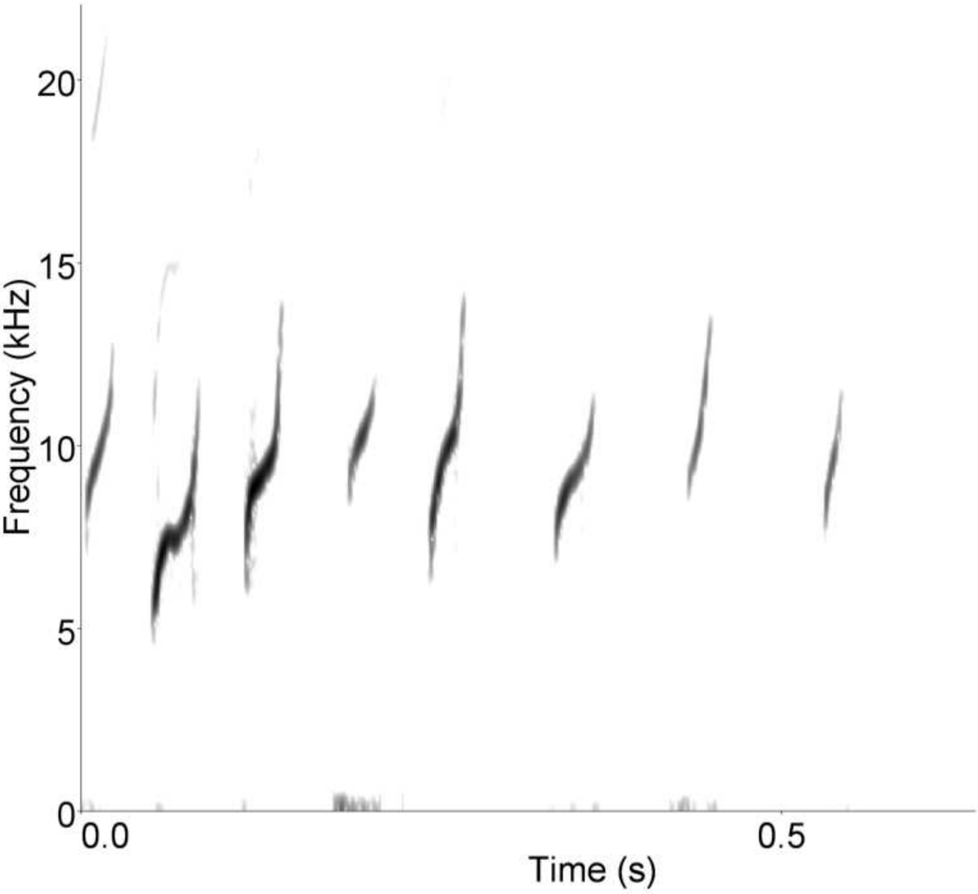
Table 2. Results of paired playback experiments on individual riflemen exposed to zip calls from first-order kin and non-kin.

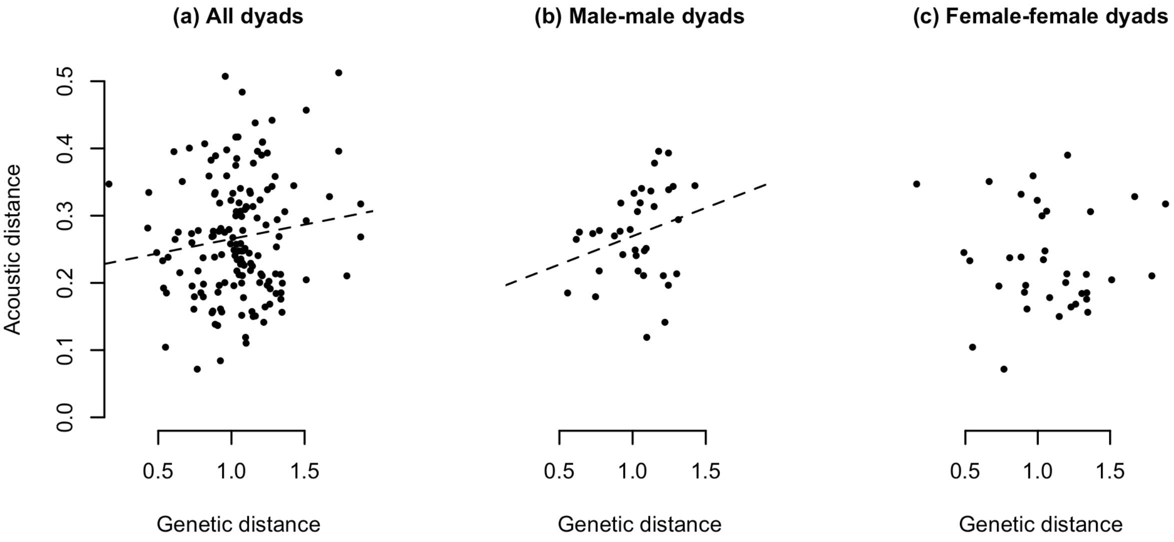
Response	**Direction**	**Test**	**Statistic**	**df**	**P**
Closest approach	Closer for kin	Paired *t*-test	$t = -1.83$	9	0.101
Approached speaker?	More likely for kin	McNemar's test	$\chi^2 = 0.57$	1	0.450
Pip calls?	Equal	McNemar's test	$\chi^2 = 0.00$	1	1.000
Zip calls?	Less likely for kin	McNemar's test	$\chi^2 = 0.00$	1	1.000
Changed behaviour?	More likely for kin	McNemar's test	$\chi^2 = 3.20$	1	0.074

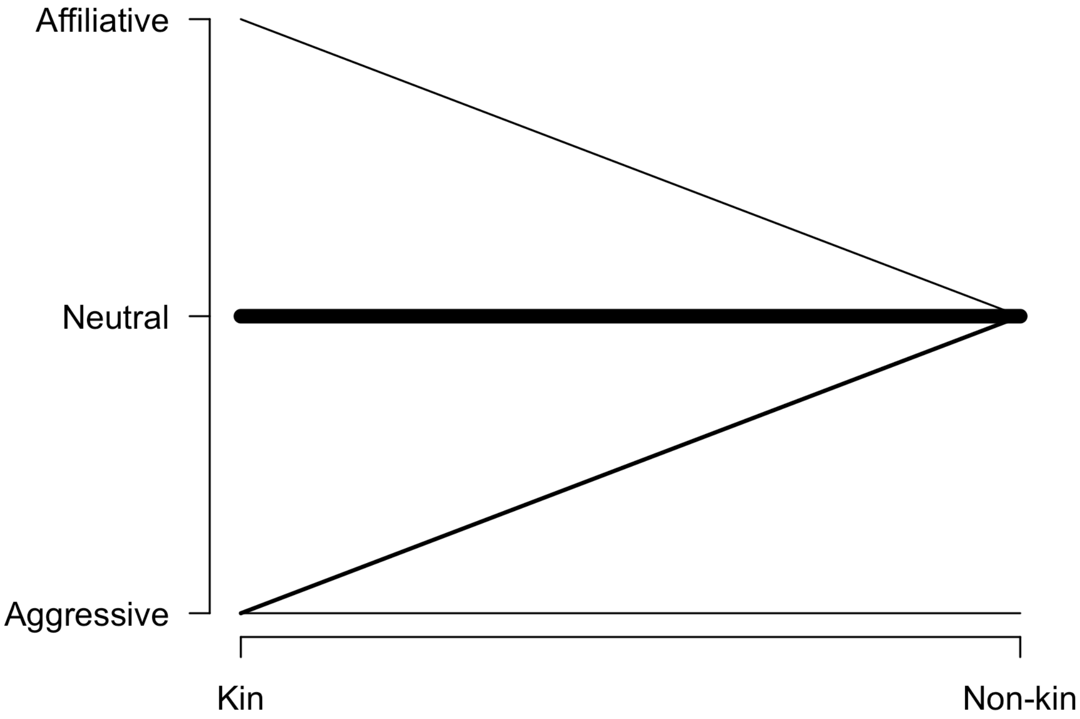
Figure 1. Example spectrograms of rifleman zip calls recorded from eight different individuals, produced using Luscinia (Lachlan et al. 2010). The leftmost shows a trace of the call's ultrasonic harmonic. Most calls show the tail at the lowest and highest frequencies, which meant that measuring minimum and maximum frequency was problematic in noisier recordings.

Figure 2. The relationship between genetic distance, measured as Queller-Goodnight relatedness subtracted from one, and difference in zip call structure (‘acoustic distance’), measured using a dynamic time warping algorithm in Luscinia (Lachlan et al. 2010), for (a) all, (b) male and (c) female dyads of adult riflemen. Trendlines illustrate statistically significant relationships.

Figure 3. Paired responses of adult riflemen to playback of the zip calls of first-order relatives and unrelated individuals. Lines connect the responses of the same individuals to respective treatments; line weights represent the number of individuals for whom the combination of responses was the same (minimum = 1, maximum = 7).


Frequency (kHz)
20
15
10
5
0
0.0
0.5
Time (s)


(a) All dyads
(b) Male-male dyads
(c) Female-female dyads
Acoustic distance
0.0
0.1
0.2
0.3
0.4
0.5
0.5
1.0
1.5
Genetic distance



583