

1 **Helping decisions and kin recognition in long-tailed tits: is call similarity used to**  
2 **direct help towards kin?**

3 Amy E. Leedale<sup>1,2</sup>, Robert F. Lachlan<sup>3</sup>, Elva J. H. Robinson<sup>4</sup> & Ben J. Hatchwell<sup>2</sup>

4 <sup>1</sup>Department of Zoology, University of Cambridge, UK

5 <sup>2</sup>Department of Animal & Plant Sciences, University of Sheffield, UK

6 <sup>3</sup>Department of Psychology, Royal Holloway University of London, UK

7 <sup>4</sup>Department of Biology, University of York, UK

8 **Abstract**

9 Most cooperative breeders live in discrete family groups, but in a minority, breeding  
10 populations comprise extended social networks of conspecifics that vary in relatedness.  
11 Selection for effective kin recognition may be expected for more related individuals in  
12 such kin neighbourhoods to maximise indirect fitness. Using a long-term social pedigree,  
13 molecular genetics, field observations and acoustic analyses, we examine how vocal  
14 similarity affects helping decisions in the long-tailed tit *Aegithalos caudatus*. Long-tailed  
15 tits are cooperative breeders in which help is typically redirected by males that have failed  
16 in their own breeding attempts towards the offspring of male relatives living within kin  
17 neighbourhoods. We identify a positive correlation between call similarity and kinship  
18 suggesting that vocal cues offer a plausible mechanism for kin discrimination.  
19 Furthermore, we show that failed breeders choose to help males with calls more similar  
20 to their own. However, although helpers fine-tune their provisioning rates according to  
21 how closely related they are to recipients, their effort was not correlated with their vocal  
22 similarity to helped breeders. We conclude that although vocalisations are an important

23 part of the recognition system of long-tailed tits, discrimination is likely to be based on  
24 prior association and may involve a combination of vocal and non-vocal cues.

25 **Keywords:** Cooperative breeding, kin recognition, kin discrimination, call similarity,  
26 acoustics.

## 27 **Introduction**

28 Kin selection is often invoked to explain the evolution of cooperation in kin-structured  
29 communities, and is expected to result in selection for some mechanism to discriminate  
30 kin from non-kin (1). A recognition mechanism that permits the differential treatment of  
31 conspecifics according to their genetic similarity (2) enables individuals to avoid  
32 inbreeding (3, 4) and maximise inclusive fitness (1, 5) in populations where kin and non-  
33 kin associate beyond reproductive maturity. However, although the adaptive functions of  
34 kin recognition are well known, the proximate mechanisms, including sensory cues and  
35 cognitive thresholds, are often difficult to determine. Our current understanding of kin  
36 recognition in social animals is that discriminating individuals acquire cues to kinship  
37 from a referent (oneself, a subset of kin, or the local environment), which are used to form  
38 internal templates (6) at a sensitive phase during development (7). Although, it may also  
39 be possible that cues and templates are genetically determined (8). Templates are later  
40 compared with the phenotypes of encountered conspecifics, and discriminatory behaviour  
41 is performed based on the perceived similarity between templates and encountered  
42 phenotypes (9). Thus, any cue that reliably co-varies with relatedness may be used to  
43 discriminate kin from non-kin.

44 Most animal societies exhibit a substantial degree of kin structure, whereby individuals  
45 are organised in to more or less discrete family units of parents and their retained offspring

46 (10-12). If the probability of encountering a relative is high, individuals can maximise  
47 indirect fitness by indiscriminately cooperating within their group (13, 14) and avoid  
48 inbreeding by selecting partners from outside the group (15). In less viscous societies,  
49 such spatial cues to kinship may be unreliable. For example, in a small number of  
50 cooperatively breeding birds, cooperation occurs after natal dispersal, across extended  
51 networks of relatives known as *kin neighbourhoods* (16). Here, the relatedness among  
52 spatially clustered individuals is less predictable, so kin recognition based on the  
53 phenotypic cues of potential social partners may be necessary (17). In such situations,  
54 selection should favour effective discrimination, but any recognition system is prone to  
55 error because phenotypic cues overlap between non-kin and kin of varying relatedness  
56 due to individual variation (6). Thus, kin recognition is likely to involve a certain rate of  
57 *acceptance errors*, where non-kin are perceived as kin, and *rejection errors*, where kin  
58 are perceived as non-kin (2, 6). The accuracy of kin recognition, and hence the frequency  
59 of such errors, depends on their relative costs, which, in turn, is determined by the  
60 probability of encountering a relative and the fitness consequences of the associated  
61 behaviours (18). This theoretical framework is supported empirically by intraspecific  
62 studies showing shifts in acceptance thresholds as the costs of error change (19), and by  
63 comparative analyses that demonstrate stronger kin discrimination in cooperatively  
64 breeding vertebrates where the benefits of helping are greater (20), and when the average  
65 relatedness within a group is lower and more variable (13).

66 Kin recognition often requires prior association; individuals learn the phenotypes of kin  
67 encountered during a sensitive phase and distinguish these familiar individuals from  
68 unfamiliar ones later in life (7). Alternatively, recognition may involve phenotype-  
69 matching, whereby individuals form a generalised template against which the phenotypes

70 of unknown individuals are compared (21). Phenotype matching does not require a period  
71 of previous association between matching individuals (22), but relies on a positive  
72 correlation between template-phenotype similarity and degree of genetic relatedness (23).  
73 Whether kin are recognised through prior association or phenotype matching can be  
74 difficult to determine; both mechanisms involve matching phenotypes to learned cues,  
75 yet they differ in template specificity (6), such that mechanisms involving phenotype  
76 matching permit individuals to recognise unfamiliar kin and distinguish between kin of  
77 varying relatedness.

78 Kin recognition based on familiarity may often be sufficient for individuals to maximise  
79 inclusive fitness by directing help towards relatives, and prior association is indeed  
80 thought to be the most common mechanism of kin recognition in cooperatively breeding  
81 birds (24-26). However, studies on long-tailed tits *Aegithalos caudatus* (27) and bell  
82 miners *Manorina melanophrys* (28), species in which helping occurs within kin  
83 neighbourhoods, found that helpers modify provisioning effort according to their degree  
84 of relatedness to recipient broods. In such situations, the risk of caring for non-kin is high,  
85 so kin recognition mechanisms with low error rates are likely to be selected for (13).  
86 Moreover, finely tuned adjustment of provisioning behaviour in relation to kinship could  
87 indicate a relatively sophisticated mechanism of kin recognition that involves phenotype  
88 matching. Vocalisations are used as kin recognition cues in both species (29, 30) and  
89 more widely in birds (31), although olfactory kin recognition has also been described in  
90 a few species (32-34). In bell miners, a relationship between genetic relatedness and vocal  
91 similarity has been reported (30), but whether this relationship exists in other species  
92 remains to be tested.

93 This study aimed to identify the mechanism permitting kin-directed cooperation and  
94 flexible helper investment in long-tailed tits, a kin-neighbourhood cooperative breeder  
95 that exhibits effective kin recognition in the absence of spatial cues (35). Helpers are  
96 failed breeders that redirect their care following unsuccessful attempts at independent  
97 breeding. Around 50% of successful nests receive help (36), typically from one or two  
98 helpers, but not all failed breeders choose to become helpers (37). Although our study  
99 population is kin-structured during breeding, most neighbours are non-kin and help is  
100 directed towards close kin more often than expected by indiscriminate helping (38).  
101 Furthermore, helpers provision more closely related broods at higher rates (27). Helpers  
102 are overwhelmingly male, and gain indirect fitness benefits by increasing the productivity  
103 of related broods (39, 40). In contrast, no direct fitness benefits of helping have been  
104 identified (41, 42). Vocalisations play a major role in in the coordination of cooperative  
105 behaviour (29). Previous studies have demonstrated individuality in the churr call: a  
106 short-range contact call often used at the nest (43). Playback and cross-fostering  
107 experiments have shown that individuals can recognize siblings using the churr call, and  
108 that these calls are learned during development (29). However, whether churr call  
109 similarity is used to assess relatedness when making helping decisions remains untested.

110 Here, we quantified variation in churr call structure within and between adult long-tailed  
111 tits and determined which sound parameters explained this variation. We also tested for  
112 an association between call similarity and relatedness, and examined whether degree of  
113 vocal similarity influenced helping decisions by analysing the churr call similarity of  
114 helpers to the breeders they helped and to nearby breeders they did not help. Finally, we  
115 investigated whether long-tailed tits helpers adjust their provisioning effort according to  
116 how similar their churr calls are to the helped breeders.

117 **Methods**

118 *Study site and field methods*

119 Fieldwork was carried out on a population of 31-46 breeding pairs of long-tailed tits in  
120 the Rivelin Valley, Sheffield, UK (53°38'N 1°56'W) from 2015 to 2017. The site is  
121 approximately 2.5km<sup>2</sup> and comprises predominantly deciduous woodland, scrub and  
122 farmland. This population of *c.*50 pairs (range: 18 – 72) has been studied extensively  
123 since 1994. The population is open: approximately 40% of breeders hatched in the study  
124 site, and are referred to as 'native' (A. E. Leedale & B. J. Hatchwell, unpublished data,  
125 2017), while the remaining 'immigrant' adults are assumed to have dispersed in to the  
126 study area during their first winter, because individuals show high breeding site fidelity  
127 following their first breeding year (40). Almost all individuals (*c.*95%) were uniquely  
128 colour-ringed for field identification. Native birds were ringed as 11-day old nestlings  
129 and immigrants were captured in mist nests and ringed under British Trust for  
130 Ornithology licence during breeding. A sample of 5-30µl of blood was taken by brachial  
131 venepuncture under Home Office licence. All nesting attempts were closely monitored to  
132 record breeding events and life-history traits such as timing of breeding, clutch size,  
133 incubation period and brood size, and the identity of parents and helpers. Nest locations  
134 were recorded using GPS receivers to an accuracy of 8m. For most nests, provisioning  
135 behaviour was observed every two days from day two of the nestling period (day 0 = day  
136 of hatching; long-tailed tit broods hatch synchronously) to fledging (typically day 16 or  
137 17) or until nest failure. Most observation periods lasted 1h, during which the identities  
138 and visit rate of all carers were recorded. For further details of provisioning observations,  
139 see (27, 39).

140 The churr call is disyllabic, consisting of an initial syllable of one or two unique elements,  
141 followed by a second syllable comprising a single element repeated several times (44).  
142 The churr calls of adult carers were recorded at the nest using a Sennheiser ME67/K6  
143 shotgun microphone onto a Roland R-05 version 1.03 WAV/MP3 recorder, with a sample  
144 rate of 48kHz, WAV-16bit accuracy, an input level of 60db and a low-cut frequency of  
145 400Hz. All recordings were made in dry conditions between 06:00 and 18:00 BST. Birds  
146 were recorded at a distance of approximately 3-15m, to minimize the effects of sound  
147 degradation and reverberation. Birds were identified by their colour ring combinations.  
148 During recording, bird ID was dictated into the microphone after each call. In total, 213  
149 recordings were made, containing 1116 churr calls from 98 individuals (mean =  $11.39 \pm$   
150  $10.24$  SD per bird; range 1 – 42).

#### 151 *Bioacoustic analysis*

152 Recordings were digitized with 16-bit accuracy at a sampling rate of 48kHz.  
153 Spectrograms were produced in Avisoft SAS-Lab Pro version 4.52 (Avisoft Bioacoustics,  
154 Raimund Specht, Berlin, Germany) using a 256-point FFT length with a Hamming  
155 window, 100% frame size and 50% window overlap, generating a frequency resolution  
156 of 188Hz and a time resolution of 2.7ms. All recordings were visualized  
157 spectrographically to assess quality. Some background noise was removed using a high-  
158 pass filter of 1.5 kHz, though recordings with extreme background noise were excluded.  
159 The sampling frequency was converted to 22.05 kHz for further analysis. As long-tailed  
160 tit calls range from 2-9kHz, this re-sampling does not affect the acoustic signal. All  
161 useable calls were stored and measured in Luscinia version 2.16.10.29.01  
162 (<https://rflachlan.github.io/Luscinia/>).

163 A subset of data was tested for individual repeatability. To determine the minimum  
164 number of calls required to capture individual variation, the cumulative repertoire size  
165 (number of distinct calls, based on number of syllables) was plotted against the number  
166 of calls considered to that point, for 100 churr calls, 10 from each of 10 birds recorded on  
167 at least two days in 2015. The resulting plots generally levelled off before the number of  
168 calls reached six (mean calls needed to reach asymptote =  $5.5 \pm 2.89$  SD, range 2-10).  
169 Therefore, repeatability tests were carried out on all calls from individuals with recordings  
170 of at least six calls from at least two days between 2015 and 2017: 907 churr calls from  
171 54 individuals (mean =  $17.46 \pm 10.02$  SD per bird; range 6 – 42). Within-individual  
172 repeatability was tested using two approaches. The first approach compared within and  
173 between-individual variation in overall call structure using dynamic time-warping  
174 (DTW), implemented in Luscinia. The second tested the individual repeatability of  
175 specific vocal characteristics (defined in Supplementary Table S1).

176 Dynamic time warping (DTW) is a distance-based programming technique used to search  
177 for an optimal alignment of two signals, which has been implemented for use in  
178 bioacoustics. The algorithm calculates a distance score between signals based on certain  
179 acoustic features, with greater distance meaning lower similarity. The acoustic features  
180 used in the DTW analysis were weighted as: time = 1, fundamental frequency = 2, change  
181 in fundamental frequency = 2, compression factor = 0.1, minimum element length = 10,  
182 time SD weighting = 1, ArcTan transform weight for frequency slope = 0.02, maximum  
183 warp = 100%. These settings generated a DTW algorithm that correctly matched visually  
184 similar vocalizations, assessed using a dendrogram and multidimensional scaling plot.  
185 This is also in line with previous studies suggesting that frequency parameters show  
186 greater individuality than temporal parameters and are particularly important for kin



187 recognition in this species (43). Pairwise comparisons of individual calls generated a  
188 matrix of DTW scores for each pair of calls. To compare call similarity within and  
189 between individuals, pairwise comparisons were assigned a value according to whether  
190 the comparison was made between calls from the same individual (0) or from two  
191 individuals (1). The DTW scores were aggregated and mean call similarity within and  
192 between individuals was compared. Since this analysis contained calls from across years,  
193 the measures of call similarity were also compared within and between years.

#### 194 *Relatedness*

195 Individuals were genotyped at 17 microsatellite loci. Genetic relatedness was estimated  
196 using coefficient of relatedness ( $r_{QG}$ ) (45) in SPAGeDi version 1.1.5 (46). This  
197 relatedness estimate is reliable when tested against our social pedigree (27). For further  
198 details on genotyping, see (47, 48). The population allele frequencies used in analyses  
199 were generated using all genotyped individuals (1994-2017,  $n = 3304$ ) in CERVUS  
200 version 3.0.7 (49) to ensure non-zero frequencies for all alleles. To calculate social  
201 relatedness among dyads, an additive relationship matrix was generated from the social  
202 pedigree (1994-2017,  $n = 3068$ ) in R version 3.5.0 (50), using the nadiv package (51). For  
203 further details on social and genetic relatedness estimates, see (38). Hereafter, genetic  
204 relatedness refers to the  $r_{QG}$  coefficients calculated from the microsatellite markers,  
205 whereas kinship refers to social relationships derived from the pedigree.

#### 206 *Call similarity, relatedness and helping*

207 Vocal similarity between individuals (1116 calls from 98 individuals) was quantified by:  
208 (i) DTW analysis, and (ii) the difference in repeatable ( $R > 0.2$ ) acoustic parameters  
209 (Table 2), measured as Euclidean distances using the R package, spaa (52). To investigate

210 how vocal similarity varied with relatedness, we tested for a relationship between churr  
211 call similarity and both genetic relatedness and kinship. For the latter, three degrees of  
212 kinship were considered: first-order ( $r = 0.5$ ), second-order ( $r = 0.25$ ), or non-kin ( $r <$   
213  $0.25$ ); non-kin relationships included only those birds for which the parentage of both  
214 birds in the dyad was known.

215 Helpers observed in 2015 – 2017 were related to male but not female parents of the broods  
216 they provisioned (see Results), so our analyses focused on helpers' vocal similarity to  
217 breeding males. If individuals use vocal similarity as a cue to relatedness, in order to  
218 direct helping effort towards close kin, helpers were expected to be more vocally similar  
219 to the breeders they helped than the breeders they did not help. For each helper, vocal  
220 similarity to male breeders at their first chosen nest in a given year ( $n = 19$ ) was compared  
221 with their mean vocal similarity to a sample of potential males (excluding those helped)  
222 nesting within 750m that year ( $n = 272$ ), the range in which the majority of failed breeders  
223 travel to provide help (mean = 337.4m  $\pm$  253.4 SD, 95% CI = 744.1m,  $n = 220$ ). Helping  
224 distance was calculated as the distance between a helper's last failed breeding attempt  
225 and the nest at which they first appeared as a helper in the same year.

226 To investigate whether helpers use vocal similarity to modify their provisioning effort,  
227 we tested for a relationship between the provisioning rates of helpers and their vocal  
228 similarity to the helped males. Because vocal similarity is a putative cue to relatedness  
229 (53), we also tested for a relationship between provisioning rate and relatedness, using  
230 genetic relatedness estimates and kinship from the social pedigree. Although the fitness  
231 consequences of helping depend on genetic relatedness, pedigree data is essential for  
232 understanding how accurately individuals are able to recognize kin, particularly when the

233 mechanism depends on socially learned cues (29). Provisioning rate was therefore  
234 expected to correlate most strongly with kinship.

### 235 *Statistical analysis*

236 Statistical analysis was carried out using R version 3.5.0 (50). Overall similarity in call  
237 structure within and between individuals was compared using a generalized linear mixed-  
238 effects model (GLMM) fitted by restricted maximum likelihood (REML) in the lme4  
239 package (54). The square root of DTW distance score was fitted as the dependent variable,  
240 comparison type (within or between individuals) as a fixed effect and bird ID 1 and bird  
241 ID 2 as nested random effects. To test for differences across years, within or between  
242 years was also fitted as a fixed effect and year 1 and year 2 fitted as nested random effects.  
243 Individual repeatability based on call parameters was carried out using multiple GLMMs  
244 in the rptR package (55, 56). Year and ID were set as grouping variables, allowing for  
245 effects of year and ID to be tested. Gaussian models were used to test the repeatability of  
246 continuous variables and Poisson models were used to test the repeatability of count  
247 variables. To test for sex differences in call characteristics, GLMMs were built with each  
248 sound parameter as the dependent variable, sex as a fixed effect and bird ID and year as  
249 random effects. The significance of fixed effects was reported using Satterthwaite's  
250 degrees of freedom in the lmerTest package (57).

251 The relationships between vocal similarity and relatedness among adult breeders were  
252 analysed using Mantel tests based on Spearman correlations of ranked distances with  
253 10000 permutations using the R package, ecodist (58). The relatedness of helpers to  
254 breeders that were helped and those that were not helped was compared using a Pearson's  
255 chi-squared test for kinship data and a general linear model fitted in lme4 for genetic

256 relatedness estimates. Vocal similarity within observed helper-breeder dyads was  
257 compared to mean vocal similarity within potential dyads for each focal helper using  
258 Wilcoxon signed-rank tests. Significance values were based on two-tailed tests and  
259 sample sizes are reported with the results. The vocal similarity of helpers to: (i) helped  
260 kin, (ii) helped non-kin, and (iii) non-kin that were not helped was compared using a  
261 GLMM fitted with DTW score measured as a continuous variable with a Gamma  
262 distribution and log link, and helper ID as a random effect.

263 To investigate whether helper provisioning rates varied with respect to their vocal  
264 similarity to male breeders, we used linear mixed-effects models fitted by REML. Genetic  
265 relatedness, kinship and vocal similarity were expected to be closely correlated, so their  
266 effect on provisioning rate was analysed in three separate models. In each model, the  
267 provisioning rate of helpers (number of visits/hour) was the response variable. In the first  
268 model, the explanatory variables were: nestling age, brood size, group size and genetic  
269 relatedness, all of which influence the provisioning rates of helpers (36). In the second  
270 model, the explanatory variables were: nestling age, brood size, group size and kinship.  
271 In the third model, the explanatory variables were: nestling age, brood size, group size  
272 and vocal similarity. Genetic relatedness was the  $r_{QG}$  estimate between helpers and male  
273 breeders, measured as a continuous variable. Kinship was the relationship between  
274 helpers and male breeders according to the social pedigree (three factor levels:  $r = 0$ ,  $r =$   
275  $0.25$  and  $r = 0.5$ ). Vocal similarity was the DTW score of churr calls between helpers and  
276 male breeders. Nestling age was measured in days from hatching (day 0). Brood size was  
277 the number of chicks in the nest on day 11, a good indicator of brood size from hatching  
278 because nestling starvation is rare (37). Group size was the number of adults that  
279 provisioned a nest (parents and helpers). Bird identity and nest identity were included as

280 random effects, to control for non-independence of repeated observations of feeding rates  
281 by the same birds, and repeated observations of feeding rates at the same nest. All  
282 explanatory covariates were initially included in full models and then dropped  
283 sequentially unless doing so significantly reduced the amount of variance explained,  
284 generating three minimum adequate models containing either genetic relatedness, kinship  
285 or vocal similarity as explanatory variables.

## 286 **Results**

### 287 *Individual repeatability*

288 Visual inspection of spectrograms suggested that churr calls from the same individual  
289 were more similar in acoustic structure than those of different individuals. This was  
290 confirmed quantitatively, because the distance score from DTW for within-individual  
291 comparisons was significantly lower than that for between-individual comparisons  
292 (GLMM, estimate  $\pm$  SE =  $-0.08 \pm 0.008$ ,  $df = 1561$ ,  $t = -9.9$ ,  $p < 0.001$ ; Fig. 1). Whether  
293 comparisons were made between calls recorded in the same or different years did not  
294 affect DTW distance score (GLMM, estimate =  $-0.003 \pm 0.005$ ,  $df = 4917$ ,  $t = -0.65$ ,  $p =$   
295  $0.55$ ). Churr calls were repeatable within individuals based on all of the parameters tested,  
296 with maximum fundamental frequency across the churr call showing the greatest  
297 individual repeatability (Table 1). There was no effect of recording year or sex on any of  
298 the parameters tested (Supplementary Table S2-S3).

### 299 *Call similarity, relatedness and helping*

300 Although there was substantial variation in vocal similarity among breeders in all  
301 pedigree categories (Supplementary Fig. S1), churr call similarity correlated with kinship

302 in several acoustic parameters: DTW score (Mantel test,  $R = -0.06$ ,  $p < 0.01$ ), difference  
303 in mean frequency ( $R = -0.04$ ,  $p = 0.03$ ) and difference in frequency bandwidth ( $R = -$   
304  $0.04$ ,  $p = 0.03$ ; Fig. 2a, Table 2). In contrast, churr call similarity did not correlate with  
305 genetic relatedness (Fig. 2b, Table 2).

306 Based on the social pedigree 32% (6/19) of helpers in 2015 – 2017 were known first-  
307 order relatives of the male, 16% (3/19) were second-order relatives of the male, and 55%  
308 (10/19) were apparently unrelated to the male. Thus, the mean relatedness of helpers to  
309 male breeders from the social pedigree was  $r = 0.19 \pm 0.2$  SD ( $n = 19$ ). The mean genetic  
310 relatedness of helpers to male breeders was  $r = 0.17 \pm 0.2$  SD ( $n = 15$ ), showing that  
311 estimates of kinship from our pedigree closely match genetically estimated relatedness.  
312 In contrast, there were no cases of help given to known female kin, and the mean genetic  
313 relatedness of helpers to females was  $r = -0.04 \pm 0.12$  SD ( $n = 13$ ). As expected, the  
314 kinship between helpers and males that they did not help within 750m was significantly  
315 lower: 10.6% ( $n = 226$ ) of relationships in which kinship was known were first order  
316 kinships, 4.4% were second order kinships, and 84.9% dyads were unrelated (Pearson's  
317 Chi-squared test,  $\chi^2 = 17.3$ ,  $p < 0.001$ ). Likewise, the mean genetic relatedness of helpers  
318 to breeding males within 750m that were not helped was  $r = 0.07 \pm 0.18$  SD ( $n = 272$ ),  
319 significantly lower than that observed for males that were helped (GLM,  $t = 2.55$ ,  $p =$   
320  $0.01$ ).

321 The decision of which male to help was positively associated with call similarity, as  
322 predicted. Failed breeders helped males that had more similar churr calls than those they  
323 did not help, based on DTW score (Wilcoxon signed-rank test:  $v = 20$ ,  $n = 19$ ,  $p < 0.01$ ;  
324 Table 3). Furthermore, there was no significant difference in the call similarity of helpers

325 to helped kin and helped non-kin (GLMM:  $t = -0.29$ ,  $n = 19$ ,  $p = 0.77$ ). In contrast, call  
326 similarity of helpers to non-kin that were helped was significantly greater than to non-kin  
327 that were not helped (GLMM:  $t = -2.52$ ,  $n = 19$ ,  $p = 0.01$ ; Fig. 3). For full model outputs,  
328 see Supplementary Table S4.

329 Finally, we examined whether helper effort was modified according to relatedness and/or  
330 call similarity, analysing the provisioning data that were available for 14 of the 19 cases  
331 of helping. These included 41 observation periods of 14 helpers at 11 nests over 3 years  
332 (mean duration of observation = 180.1 min  $\pm$  145.3 SD per nest, range = 1 – 8h, mean  
333 feeding rate (visits per hour) = 5.05  $\pm$  2.56 SD per helper, range = 1 – 10.4). The  
334 provisioning rate of helpers increased with kinship between helper and male breeder  
335 (GLMM:  $\chi^2 = 5.68$ ,  $p = 0.02$ ), an effect that increased with group size (GLMM:  $\chi^2 = 11.61$ ,  
336  $p < 0.001$ ). The provisioning rate of helpers also increased with genetic relatedness  
337 between helper and male breeder when group size was large (GLMM:  $\chi^2 = 5.94$ ,  $p = 0.01$ ).  
338 In contrast, the vocal similarity between helper and male breeder had no effect on helper  
339 provisioning rate (GLMM:  $\chi^2 = 0.01$ ,  $p = 0.9$ ). For full model comparisons, see  
340 Supplementary Table S5.

## 341 **Discussion**

342 Vocalisations are clearly a critical component of the kin-selected cooperative breeding  
343 system of long-tailed tits. Our results show, along with previous studies, that these calls  
344 are individual-specific (43) and that churr call similarity is positively associated with  
345 kinship (53), although this was the case for the social pedigree but not for genetic  
346 relatedness estimated from microsatellite data. This finding is consistent with previous  
347 experimental studies indicating that churr calls are learned from provisioning adults in

348 early development (29). The sample of helpers included in this study showed a strong  
349 preference for kin relative to their availability, a finding that is again consistent with  
350 previous observational and experimental studies (35, 38). Importantly, we found that  
351 churr calls offer a potential mechanism to facilitate this kin preference because the calls  
352 of helpers were more similar to males they helped than to those they did not. Moreover,  
353 some helpers assisted unrelated recipients, and we found that call similarity between  
354 helpers and these non-kin recipients was greater than that between helpers and non-kin  
355 they could have helped. By contrast, there was no difference in the call similarity of  
356 helpers to kin and non-kin recipients. Finally, contrary to expectations, we found that  
357 although the provisioning effort of helpers was correlated with kinship, again supporting  
358 the findings of a previous study (27), helper effort was not predicted by call similarity to  
359 the helped male breeder.

360 Previous studies have revealed a strong preference for kin by helpers in long-tailed tits  
361 (35, 38), as reported in many other cooperatively breeding vertebrates (e.g. 13, 20, 24,  
362 30, 59). Studies of other species have also shown that kin recognition is achieved using  
363 vocal cues (e.g. 60-62), and this study provides further direct evidence that vocal  
364 similarity is the mechanism of kin recognition that permits kin-directed helping in long-  
365 tailed tits. However, there remains the persistent puzzle that a significant proportion of  
366 helpers in this species help broods to which they are unrelated (27, 38, 41), even though  
367 they appear to gain no benefit from doing so (42). As vocalisations are learned in the nest,  
368 it is possible that helpers (whether related or unrelated) could gain future direct benefits  
369 through increasing the chances of being identified as kin by the grown offspring of the  
370 broods they helped. However, reciprocal helping is rare; in just 3.7% cases did helpers  
371 choose to help breeders that had helped at their natal nest. Indeed, the high annual



372 mortality rate provides little opportunity for reciprocity from helped broods (42), and  
373 most helping occurs among siblings (27, 35). It therefore seems unlikely that the  
374 opportunity to be identified as kin by helped broods could drive helper decisions. Instead,  
375 our results support an earlier suggestion that this counter-intuitive behaviour arises from  
376 recognition errors (41).

377 The theoretical framework of the acceptance threshold model argues that an actor  
378 categorises conspecifics depending on an acceptance threshold: a degree of template-  
379 phenotype dissimilarity below which it will accept and above which it will reject  
380 conspecifics as kin (6). Our results suggest that long-tailed tits use degree of vocal  
381 similarity to recognise close kin, but also that their threshold for kin discrimination does  
382 lead to acceptance errors, with non-kin sometimes helped despite there being no known  
383 benefit of doing so (38, 41, 42). There are two reasons why long-tailed tit helpers may be  
384 'generous' with their help and inclined to make acceptance errors. First, although failed  
385 breeders may prefer to help close kin, given that relatives are clustered within the range  
386 that most helping occurs (38), there might still be a reasonable chance of gaining some  
387 indirect fitness by helping an unfamiliar individual because they could be more distantly  
388 related. This suggestion is supported by the finding of Leedale *et al.* (38) that the  
389 frequency of helping second order relatives was as expected from random choice among  
390 nearby males. Secondly, Hatchwell *et al.* (41) argued that the costs of helping are low in  
391 long-tailed tits because there is no cost of lost breeding opportunity (all helpers are  
392 breeders that have failed to reproduce successfully) and help is provided for only a short  
393 period during the nestling and post-fledging stage. In contrast, the potential benefit of  
394 helping, via the increased recruitment of relatives is high. Therefore, a permissive  
395 threshold for acceptance of another individual as kin should be selected for (6). The

396 critical finding from that previous study (41) is that even with low mean relatedness  
397 between helpers and the brood they care for ( $r = 0.17$ ), Hamilton's rule for the evolution  
398 of altruistic helping was satisfied.

399 Our finding that social pedigree was a better predictor of vocal similarity than genetic  
400 relatedness estimates was expected given that churr calls are learned (29). Several other  
401 species of cooperatively breeding birds have family- or group-specific vocalisations that  
402 are also consistent with this mechanism (60, 63-66). Learned kin recognition cues in any  
403 sensory modality are expected to be reliable if they are acquired at a time when  
404 associating individuals are close kin. In long-tailed tits, churr calls develop in the nest,  
405 learned from tutors that are likely to be first-order relatives, increasing vocal similarity  
406 among first-order kin relative to the general population (29). Nevertheless, a continuous  
407 positive correlation between call similarity and relatedness, rather than a threshold effect,  
408 could arise if calls are learned from parents; for example, half-siblings that share one  
409 parent could be less vocally similar than full siblings that share two parents, or uncles,  
410 aunts and even cousins could conceivably retain some family-specific vocal traits.  
411 Although genetic relatedness estimates are reliable when tested against our social  
412 pedigree (27), the variation and overlap in genetic relatedness estimates for first-order,  
413 second-order and non-kin (Supplementary Fig. S1) may explain the weak correlation  
414 between vocal similarity and genetic relatedness estimates compared with pedigree  
415 kinship (53).

416 Our results do not exclude a genetic influence on vocal variation; indeed, high individual  
417 repeatability suggests some innate individual differences in long-tailed tits. Yet, any  
418 recognition system that relies entirely on genetically acquired cues may be susceptible to  
419 rejection errors because mutation and recombination would cause even close kin to be

420 genetically dissimilar at some loci (67, 68). Genetic recognition cues have been reported  
421 in several non-avian taxa (69-70), but in social birds, kin recognition typically requires a  
422 critical period of learning during which recognition templates are formed (7). However,  
423 very little is known about how socially learned recognition cues develop; for example,  
424 which adults act as tutors, or whether offspring can distinguish between helpers and  
425 parents during learning is unknown in any cooperative breeder. Further investigation into  
426 the learning and development of vocal cues in situations where there are multiple potential  
427 tutors is a worthwhile avenue for further study.

428 Familiarity is the most widely supported mechanism of kin recognition in cooperatively  
429 breeding birds (71, 72), with kin association during extended brood care providing the  
430 sensitive period during which reliable recognition templates can form (7). In long-tailed  
431 tits, it is very likely that first-order kin are associated during this crucial period, but there  
432 are instances in which this is not the case. First, extra-pair paternity occurs at low rates  
433 and results in half-siblings being raised together (73). Second, offspring presumably  
434 acquire recognition templates from any second-order kin or non-kin that helped them as  
435 a nestling. Third, pair-bonds that last more than one year may produce full-siblings that  
436 have not been raised together, although the high annual mortality rate (42), high divorce  
437 rate (74) and low probability of successful reproduction (39) together make this a rare  
438 event. But, despite these complicating factors, the pattern that long-tailed tits usually help  
439 at the nest of individuals with whom they have had close prior association (29, 75)  
440 supports the idea that familiarity is the principal driver of helping decisions. Familiarity  
441 is also a stronger predictor of cooperative behaviour than genetic relatedness in Galápagos  
442 mockingbirds *Nesomimus parvulus* (24) and Seychelles warblers *Acrocephalus*  
443 *sechellensis* (26). In the latter species, helpers provision the offspring of breeding females

444 that provisioned them as a nestling, suggesting the context of prior association influences  
445 helper decisions (76).

446 Although kinship to male breeders explained a considerable amount of the variation in  
447 the provisioning rates of individual helpers, helpers did not adjust their provisioning rates  
448 according to vocal similarity to those breeders, suggesting that churr call similarity alone  
449 is not responsible for the fine-tuning of helping effort in relation to kinship. Therefore,  
450 although vocalisations may convey kinship cues, assessment of relatedness based solely  
451 on degree of call similarity is unlikely. This contrasts with studies of the bell miner, which  
452 identified ‘mew’ call similarity as the cue to relatedness that allows helpers to make fine-  
453 scale adjustments in their provisioning effort (28, 30). However, whether the reported  
454 relationship between provisioning effort and call similarity in bell miners is continuous  
455 or threshold-based is unclear. Bell miners live in complex societies in which membership  
456 of a coterie does not guarantee kinship and there is no evidence of a period of call learning,  
457 suggesting that ‘mew’ calls are innate (30). Such genetically acquired cues would permit  
458 bell miners to recognise relatives in a population where familiarity does not signal  
459 kinship. By contrast, the social structure of long-tailed tits is relatively simple. The  
460 proximity of non-kin and kin of variable relatedness within breeding populations also  
461 necessitates active kin recognition, but the period of more or less exclusive association  
462 between close kin in early life provides an opportunity to learn the identity of kin that is  
463 missing in bell miners.

## 464 **Conclusions**

465 Our results indicate that vocal similarity is part of a combination of cues that allows  
466 individual long-tailed tits to recognise familiar individuals. The positive relationship

467 between provisioning effort and relatedness to the brood may reflect a decision to help  
468 more familiar kin at a higher rate than less familiar individuals that are likely to be more  
469 distant kin. Discrimination based on prior association or familiarity would permit this  
470 adjustment. Regarding the precise cues used for discrimination, whether they are learned  
471 or genetic, a recognition mechanism that depends on variation in a single trait may be  
472 unstable; individuals bearing common cues are more likely to be accepted as kin than  
473 those with rare cues, leading to phenotypic convergence or fixation, and subsequent  
474 breakdown of the recognition system (77). A recognition system based on multi-  
475 component kin 'signatures' would be less vulnerable to such processes.

#### 476 **Acknowledgements**

477 We are grateful to all those who have contributed to the long-tailed tit project, and thank  
478 René van Dijk and Stuart Sharp for discussion. Molecular analyses were conducted at the  
479 NERC Biomolecular Analysis Facility at the University of Sheffield, with support from  
480 Terry Burke, Deborah Dawson, Natalie dos Remedios and Maria-Elena Mannarelli. This  
481 work was funded by the Natural Environment Research Council (NERC, UK, 1517208  
482 and NE/I027118/1).

#### 483 **References**

- 484 1. Hamilton W. D. (1964). The genetical evolution of social behaviour (I and II). *J.*  
485 *Theor. Biol.*, **7**, 1-52.
- 486 2. Sherman P. W., Reeve H. K. & Pfennig D. W. (1997). Recognition Systems. In:  
487 *Behavioral Ecology: An Evolutionary Approach* (eds. J. R. Krebs & N. B. Davies),  
488 pp. 69-96. Cambridge, UK: Blackwell Science Ltd.
- 489 3. Pusey A. E. & Wolf M. (1996). Inbreeding avoidance in animals. *Trends Ecol.*  
490 *Evol.*, **11**, 201-206.

- 491 4. Koenig W. D. & Haydock J. L. (2004). Incest and incest avoidance. In: *Ecology*  
492 *and Evolution of Cooperative Breeding in Birds* (eds. W. D. Koenig & J. L.  
493 Dickinson), pp. 142-156. Cambridge, UK: Cambridge University Press.
- 494 5. Bourke A. F. G. (2011). *Principles of Social Evolution*. Oxford Series in Ecology  
495 and Evolution (eds. P. H. Harvey, R. M. May, C. H. Godfray & J. A. Dunne).  
496 Oxford: Oxford University Press.
- 497 6. Reeve H. K. (1989). The evolution of conspecific acceptance thresholds. *Am. Nat.*,  
498 **133**, 407-435.
- 499 7. Komdeur J. & Hatchwell B. J. (1999). Kin recognition: function and mechanism  
500 in avian societies. *Trends Ecol. Evol.*, **14**, 237-241.
- 501 8. Queller D. C., Ponte E., Bozzaro S. & Strassmann J. E. (2003). Single-Gene  
502 Greenbeard Effects in the Social Amoeba *Dictyostelium discoideum*. *Science*,  
503 **299**, 105-106.
- 504 9. Lacy R. C. & Sherman P. W. (1983). Kin recognition by phenotypic matching.  
505 *Am. Nat.*, **121**, 489-512.
- 506 10. Hatchwell B. J. (2009). The evolution of cooperative breeding in birds: kinship,  
507 dispersal and life history. *Philos. T. Roy. Soc. B*, **364**, 3217–3227.
- 508 11. Riehl C. (2013). Evolutionary routes to non-kin cooperative breeding in birds.  
509 *Proc. Roy. Soc. B*, **280**, 20132245–20132245.
- 510 12. Rubenstein D. R. & Abbot P. (2017). *Comparative Social Evolution*. Cambridge,  
511 UK: Cambridge University Press.
- 512 13. Cornwallis C. K., West S. A. & Griffin A. S. (2009). Routes to indirect fitness in  
513 cooperatively breeding vertebrates: kin discrimination and limited dispersal. *J.*  
514 *Evol. Biol.*, **22**, 2445-2457.
- 515 14. Duncan C., Gaynor D., Clutton-Brock T. & Dyble M. (2019). The Evolution of  
516 Indiscriminate Altruism in a Cooperatively Breeding Mammal. *Am. Nat.*, **193**,  
517 841-851.

- 518 15. Varian-Ramos C. W. & Webster M. S. (2012). Extrapair copulations reduce  
519 inbreeding for female red-backed fairy-wrens, *Malurus melanocephalus*. *Anim.*  
520 *Behav.*, **83**, 857-864.
- 521 16. Dickinson J. L. & Hatchwell B. J. (2004). Fitness consequences of helping. In:  
522 *Ecology and Evolution of Cooperative Breeding in Birds* (eds. W. D. Koenig & J.  
523 L. Dickinson), pp 48-66. Cambridge, UK: Cambridge University Press.
- 524 17. Gamboa G. J., Reeve H. K. & Holmes W. G. (1991). Conceptual issues and  
525 methodology in kin-recognition research - a critical discussion. *Ethology*, **88**, 109-  
526 127.
- 527 18. Agrawal A. (2001). Kin recognition and the evolution of altruism. *Proc. Roy. Soc.*  
528 *B*, **268**, 1099-1104.
- 529 19. Downs S. G. & Ratnieks F. L. W. (2000). Adaptive shifts in honey bee (*Apis*  
530 *mellifera* L.) guarding behavior support predictions of the acceptance threshold  
531 model. *Behav. Ecol.*, **11**, 326-333.
- 532 20. Griffin A. S. & West S. A. (2003). Kin discrimination and the benefit of helping  
533 in cooperatively breeding vertebrates. *Science*, **302**, 634-636.
- 534 21. Holmes W. G. & Sherman P. W. (1983). Kin recognition in animals. *Am. Sci.*, **71**,  
535 46-55.
- 536 22. Tang-Martinez Z. (2001). The mechanisms of kin discrimination and the  
537 evolution of kin recognition in vertebrates: a critical re-evaluation. *Behav.*  
538 *Process.*, **53**, 21-40.
- 539 23. Mateo J. M. (2004). Recognition systems and biological organization: The  
540 perception component of social recognition. *Ann. Zool. Fenn.*, **41**, 729-745.
- 541 24. Curry R. L. (1988). Influence of kinship on helping behaviour of Galápagos  
542 Mockingbirds. *Behav. Ecol. Sociobiol.*, **22**, 141-152.
- 543 25. Hatchwell B. J., Ross D. J., Fowlie M. K. & McGowan A. (2001). Kin  
544 discrimination in cooperatively breeding long-tailed tits. *Proc. Roy. Soc. B*, **268**,  
545 885-890.

- 546 26. Komdeur J., Richardson D. S. & Burke T. (2004). Experimental evidence that kin  
547 discrimination in the Seychelles warbler is based on association and not on genetic  
548 relatedness. *Proc. Roy. Soc. B*, **271**, 963-969.
- 549 27. Nam K.-B., Simeoni M., Sharp S. P. & Hatchwell B. J. (2010). Kinship affects  
550 investment by helpers in a cooperatively breeding bird. *Proc. Roy. Soc. B* **277**,  
551 3299-3306.
- 552 28. Wright J., McDonald P. G., te Marvelde L., Kazem A. J. N. & Bishop C. M.  
553 (2010). Helping effort increases with relatedness in bell miners, but ‘unrelated’  
554 helpers of both sexes still provide substantial care. *Proc. Roy. Soc. B*, **227**, 437-  
555 445.
- 556 29. Sharp S. P., McGowan A., Wood M. J. & Hatchwell B. J. (2005). Learned kin  
557 recognition cues in a social bird. *Nature*, **434**, 1127-1130.
- 558 30. McDonald P. G. & Wright J. (2011). Bell miner provisioning calls are more  
559 similar among relatives and are used by helpers at the nest to bias their effort  
560 towards kin. *Proc. Roy. Soc. B*, **278**, 3403-3411.
- 561 31. Bradbury, J. W. & Vehrencamp S. L. (1998). *Principles of Animal*  
562 *Communication*. Sinauer, Sunderland.
- 563 32. Coffin H. R., Watters J. V. & Mateo J. M. (2011). Odour-Based Recognition of  
564 Familiar and Related Conspecifics: A First Test Conducted on Captive Humboldt  
565 Penguins (*Spheniscus humboldti*). *Plos One*, 6, 1-4.
- 566 33. Krause E. T., Kruger O., Kohlmeier P. & Caspers B. A. (2012). Olfactory kin  
567 recognition in a songbird. *Biol. Letters*, **8**, 327-329.
- 568 34. Bonadonna F. & Sanz-Aguilar A. (2012). Kin recognition and inbreeding  
569 avoidance in wild birds: the first evidence for individual kin-related odour  
570 recognition. *Anim. Behav.*, **84**, 509-513.
- 571 35. Russell A. F. & Hatchwell B. J. (2001). Experimental evidence for kin-biased  
572 helping in a cooperatively breeding vertebrate. *Proc. Roy. Soc. B*, **268**, 2169-2174.



- 573 36. Hatchwell B. J., Sharp S. P., Beckerman A. P. & Meade J. (2013). Ecological and  
574 demographic correlates of helping behaviour in a cooperatively breeding bird. *J.*  
575 *Anim. Ecol.*, **82**, 486-494.
- 576 37. MacColl A. D. C. & Hatchwell B. J. (2004). Determinants of lifetime fitness in a  
577 cooperative breeder, the long-tailed tit *Aegithalos caudatus*. *J. Anim. Ecol.*, **73**,  
578 1137–1148.
- 579 38. Leedale A. E., Sharp S. P., Simeoni M., Robinson E. J. H. & Hatchwell B. J.  
580 (2018). Fine-scale genetic structure and helping decisions in a cooperatively  
581 breeding bird. *Mol. Ecol.*, **27**, 1714–1726.
- 582 39. Hatchwell B. J., Russell A. F., MacColl A. D. C., Ross D. J., Fowlie M. K. &  
583 McGowan A. (2004). Helpers increase long-term but not short-term productivity  
584 in cooperatively breeding long-tailed tits. *Behav. Ecol.*, **15**, 1–10.
- 585 40. McGowan A., Hatchwell B. J. & Woodburn R. J. W. (2003). The effect of helping  
586 behaviour on the survival of juvenile and adult long-tailed tits *Aegithalos*  
587 *caudatus*. *J. Anim. Ecol.*, **72**, 491-499.
- 588 41. Hatchwell B. J., Gullett P. R. & Adams M. J. (2014). Helping in cooperatively  
589 breeding long-tailed tits: a test of Hamilton’s rule. *Philos. T. Roy. Soc. B*, **369**,  
590 20130565.
- 591 42. Meade J. & Hatchwell B. J. (2010). No direct fitness benefits of helping in a  
592 cooperative breeder despite higher survival of helpers. *Behav. Ecol.* **21**, 1186-  
593 1194.
- 594 43. Sharp S. P. & Hatchwell B. J. (2006). Development of family specific contact  
595 calls in the long-tailed tit *Aegithalos caudatus*. *Ibis*, **148**, 649-656.
- 596 44. Sharp S. P. & Hatchwell B. J. (2005). Individuality in the contact calls of  
597 cooperatively breeding long-tailed tits (*Aegithalos caudatus*). *Behaviour*, **142**,  
598 1559-1575.
- 599 45. Queller D. C. & Goodnight K. F. (1989). Estimating relatedness using genetic  
600 markers. *Evolution*, **43**, 258-275.

- 601 46. Hardy O. J. & Vekemans X. (2002). SPAGeDi: a versatile computer program to  
602 analyse spatial genetic structure at the individual or population levels. *Mol. Ecol.*  
603 *Notes*, **2**, 618–620.
- 604 47. Simeoni M., Dawson D. A., Ross D. J., Châline N., Burke T. & Hatchwell B. J.  
605 (2007) Characterization of 20 microsatellite loci in the long-tailed tit *Aegithalos*  
606 *caudatus* (Aegithalidae, AVES). *Mol. Ecol. Notes*, **7**, 1319-1322.
- 607 48. Adams M. J., Robinson M. R., Mannarelli M-E. & Hatchwell B. J. (2015). Social  
608 genetic and social environment effects on parental and helper care in a  
609 cooperatively breeding bird. *Proc. Roy. Soc. B*, **282**, 20150689.
- 610 49. Kalinowski S. T., Taper M. L. & Marshall T. C. (2007). Revising how the  
611 computer program CERVUS accommodates genotyping error increases success  
612 in paternity assignment. *Mol. Ecol.*, **16**, 1099-1106.
- 613 50. R Core Team (2018). *R: A language and environment for statistical computing*. R  
614 Foundation for Statistical Computing, Vienna, Austria. URL [https://www.R-](https://www.R-project.org/)  
615 [project.org/](https://www.R-project.org/).
- 616 51. Wolak M. E. (2012). nadiv: an R package to create relatedness matrices for  
617 estimating non-additive genetic variances in animal models. *Methods Ecol. Evol.*,  
618 **3**, 792-796.
- 619 52. Zhang J. (2016). *spaa: SPecies Association Analysis*. R package version 0.2.1,  
620 URL <https://CRAN.R-project.org/package=spaa>.
- 621 53. Leedale A. E. (2018). Functions and mechanisms of kin recognition in long-tailed  
622 tits. *PhD Thesis*.
- 623 54. Bates D., Maechler M., Bolker B. & Walker S. (2015). Fitting Linear Mixed-  
624 Effects Models Using *lme4*. *J. Stat. Softw.*, **67**, 1-48.
- 625 55. Nakagawa S. & Schielzeth H. (2010). Repeatability for Gaussian and non-  
626 Gaussian data: a practical guide for biologists. *Biol. Rev.*, **85**, 935-956.
- 627 56. Schielzeth H., Nakagawa S. & Stoffel M. (2011). *rptR: Repeatability for Gaussian*  
628 *and non-Gaussian data*. R package version 0.6.405.

- 629 57. Kuznetsova A., Brockhoff P B. & Christensen R. H. B. (2017). *lmerTest* Package:  
630 Tests in Linear Mixed Effects Models. *J Stat. Softw.*, **82**, 1–26.
- 631 58. Goslee S. & Urban D. (2007). The ECODIST package for dissimilarity-based  
632 analysis of ecological data. *J Stat. Softw.*, **22**, 1-19. 10.18637/jss.v022.i07.
- 633 59. Emlen S. T. & Wrege P. H. (1988). The role of kinship in helping decisions among  
634 white-fronted bee-eaters. *Behav. Ecol. Sociobiol.*, **23**, 305-315.
- 635 60. Hopp S. L., Jablonski P. & Brown J. L. (2001). Recognition of group membership  
636 by voice in Mexican jays, *Aphelocoma ultramarina*. *Anim. Behav.*, **62**, 297–303.
- 637 61. Payne R. B., Payne L. L. & Rowley I. (1988). Kin and social relationships in  
638 splendid fairy-wrens - recognition by song in a cooperative bird. *Anim. Behav.*,  
639 **36**, 1341-1351.
- 640 62. Price J. J. (1998). Family- and sex-specific vocal traditions in a cooperatively  
641 breeding songbird. *Proc. Roy. Soc. B*, **265**, 497-502.
- 642 63. Akçay Ç., Hambury K. L., Arnold J. A., Nevins A. M. & Dickinson J. L. (2014).  
643 Song sharing with neighbours and relatives in a cooperatively breeding songbird.  
644 *Anim. Behav.*, **92**, 55-62.
- 645 64. Akçay Ç., Swift R. J., Reed V. A. & Dickinson J. L. (2013). Vocal kin recognition  
646 in kin neighborhoods of western bluebirds. *Behav. Ecol.*, **24**, 898-905.
- 647 65. Radford A. N. (2005). Group-specific vocal signatures and neighbour-stranger  
648 discrimination in the cooperatively breeding green woodhoopoe. *Anim. Behav.*,  
649 **70**, 1227-1234.
- 650 66. Crane JMS, JL Pick, AJ Tribe, E Vincze, BJ Hatchwell & AF Russell (2015)  
651 Chestnut-crowned babbblers show affinity for calls of removed group members: a  
652 dual playback without expectancy violation. *Anim. Behav.*, **104**, 51-57.
- 653 67. Dawkins R. (1976). *The Selfish Gene*. Oxford, UK: Oxford University Press.
- 654 68. Keller L. & Ross K. G. (1998). Selfish genes: a green beard in the red fire ant.  
655 *Nature*, **394**, 573-575.

- 656 69. Gamboa G. J., Reeve H. K., Ferguson I. D. & Wacker T. L. (1986). Nestmate  
657 recognition in social wasps - the origin and acquisition of recognition odours.  
658 *Anim. Behav.*, **34**, 685-695.
- 659 70. Green J. P., Holmes A. M., Davidson A. J., Paterson S., Stockley P., Beynon R.  
660 J., & Hurst J. L. (2015). The genetic basis of kin recognition in a cooperatively  
661 breeding mammal. *Curr. Biol.*, **25**, 2631-2641.
- 662 71. Riehl C. & Stern C. A. (2015). How cooperatively breeding birds identify relatives  
663 and avoid incest: New insights into dispersal and kin recognition. *BioEssays*, **37**,  
664 1303–1308.
- 665 72. Leedale A. E., Li J., Hatchwell B. J. (2020). Kith or Kin? Familiarity as a Cue to  
666 Kinship in Social Birds. *Frontiers in Ecology and Evolution* **8**, 77.
- 667 73. Hatchwell B. J., Ross D. J., Chaline N., Fowlie M. K., Burke T. (2002). Parentage  
668 in the cooperative breeding system of long-tailed tits, *Aegithalos caudatus*. *Anim.*  
669 *Behav.*, **64**, 55-63.
- 670 74. Hatchwell B. J., Russell A. F., Ross D. J. & Fowlie M. K. (2000). Divorce in  
671 cooperatively breeding long-tailed tits: a consequence of inbreeding avoidance?  
672 *Proc. Roy. Soc. B*, **267**, 813-819.
- 673 75. Napper C. J. & Hatchwell B. J. (2016). Social dynamics in nonbreeding flocks of  
674 a cooperatively breeding bird: causes and consequences of kin associations. *Anim.*  
675 *Behav.*, **122**, 23-35.
- 676 76. Richardson D. S., Burke T. & Komdeur J. (2003). Sex-specific associative  
677 learning cues and inclusive fitness benefits in the Seychelles warbler. *J. Evol.*  
678 *Biol.*, **16**, 854-861.
- 679 77. Crozier R. H. (1986). Genetic clonal recognition abilities in marine invertebrates  
680 must be maintained by selection for something else. *Evolution*, **40**, 1100-1101.

681 **Table 1.** Repeatability of long-tailed tit churr call parameters based on caller identity (*n*  
 682 = 907 calls from 54 birds).

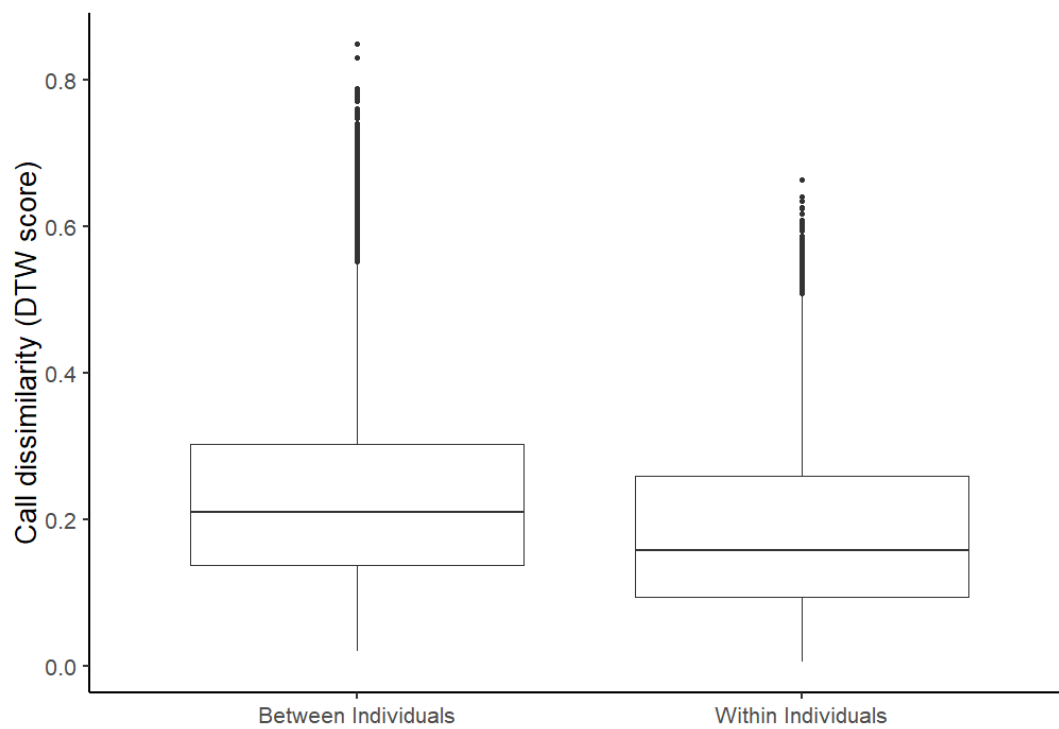
Call parameter	<i>R</i> ± SE	CI	<i>p</i>
Duration (ms)	0.33 ± 0.05	0.22, 0.42	< 0.001
Number of repeats	0.07 ± 0.02	0.02, 0.09	< 0.001
Fundamental frequency (Hz)	0.29 ± 0.05	0.19, 0.38	< 0.001
Maximum fundamental frequency (Hz)	0.5 ± 0.07	0.35, 0.61	< 0.001
Bandwidth (Hz)	0.21 ± 0.05	0.13, 0.3	< 0.001
Weiner Entropy	0.19 ± 0.04	0.11, 0.27	< 0.001

683 **Table 2.** The correlation between churr call dissimilarity and relatedness in long-tailed  
684 tits based on dynamic time-warping analysis (DTW) and the difference ( $\Delta$ ) in a range of  
685 acoustic parameters. The results and significance values from Mantel tests are reported  
686 for dyadic comparisons among breeders based on degree of kinship calculated from the  
687 social pedigree ( $n = 80$ ), and genetic relatedness estimates ( $n = 88$ ).

Measure of call dissimilarity	Relatedness variable	Mantel $R$	Lower 5% CI	Upper 95% CI	$p$
DTW	Kinship	-0.06	-0.08	-0.05	0.001
$\Delta$ Bandwidth (Hz)	Kinship	-0.04	-0.05	-0.03	0.028
$\Delta$ Mean Frequency (Hz)	Kinship	-0.04	-0.05	-0.02	0.029
$\Delta$ Max. Frequency (Hz)	Kinship	-0.01	-0.02	0.01	0.500
$\Delta$ Duration (ms)	Kinship	-0.03	-0.04	-0.01	0.116
DTW	Genetic	-0.01	-0.02	0.01	0.819
$\Delta$ Bandwidth (Hz)	Genetic	-0.03	-0.04	-0.01	0.281
$\Delta$ Mean Frequency (Hz)	Genetic	0.01	-0.01	0.02	0.661
$\Delta$ Max. Frequency (Hz)	Genetic	0.02	0.01	0.04	0.278
$\Delta$ Duration (ms)	Genetic	-0.03	-0.05	-0.01	0.227

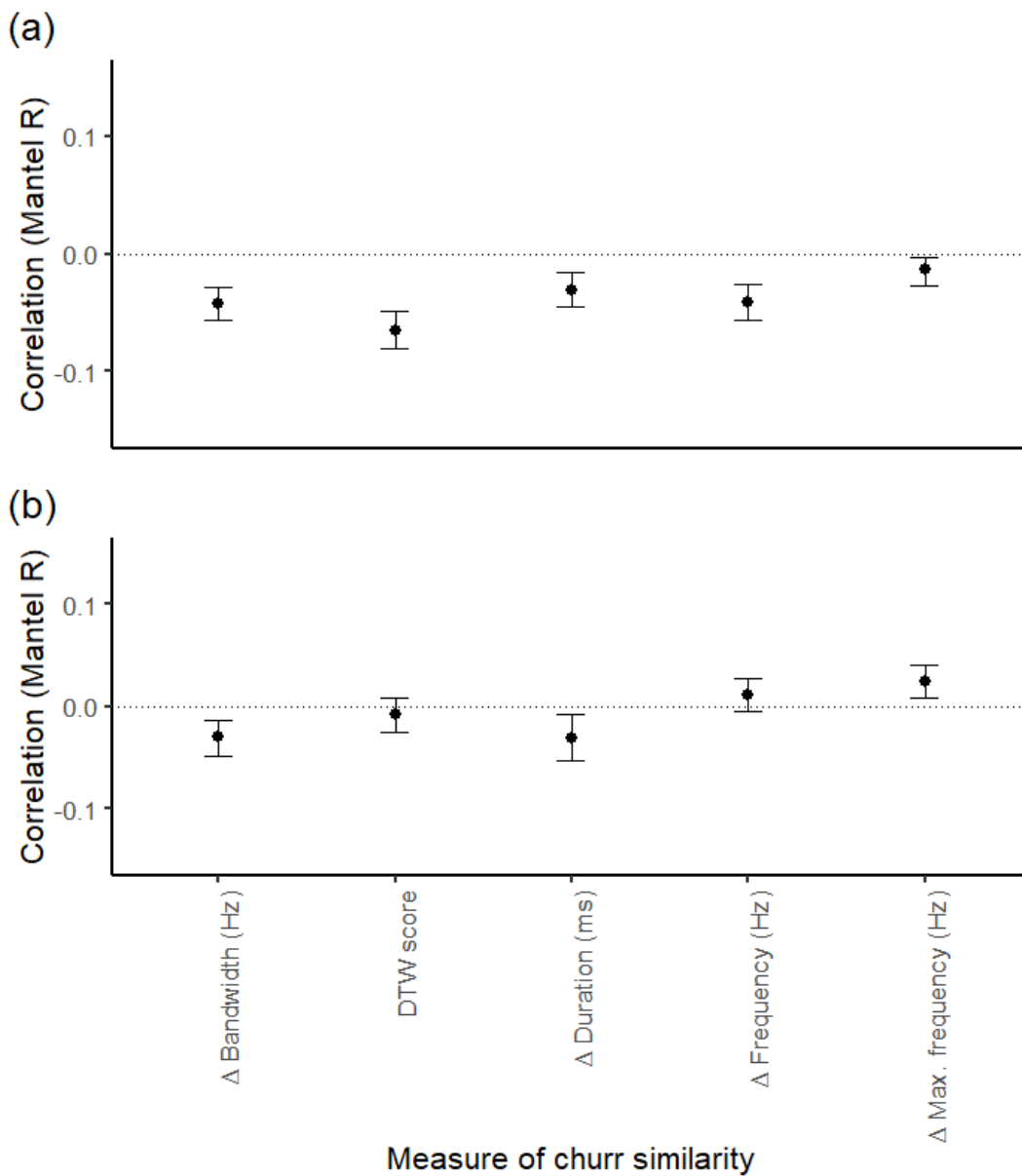
688 **Table 3.** Dissimilarity of churr calls between helpers and helped male breeders ( $n = 19$ )  
689 compared with the mean call dissimilarity of those helpers to the potential males they did  
690 not help ( $n = 272$ ). Potential males were those breeding within 750m of the helpers last  
691 failed nest in the same year. Call dissimilarity was measured using dynamic time warping  
692 analysis (DTW), and the difference ( $\Delta$ ) in a range of acoustic parameters.

Measure of call dissimilarity	Helped males ( $n = 19$ dyads)	Potential males ( $n = 252$ dyads)	Wilcoxon's signed rank	
	mean $\pm$ SD	mean $\pm$ SD	$\nu$	$p$
DTW score	0.21 $\pm$ 0.06	0.26 $\pm$ 0.06	20	0.002
$\Delta$ Bandwidth (Hz)	149.54 $\pm$ 138.48	193.93 $\pm$ 103.52	61	0.18
$\Delta$ Mean frequency (Hz)	274.07 $\pm$ 207.49	318.93 $\pm$ 117.59	63	0.21
$\Delta$ Maximum frequency (Hz)	243.61 $\pm$ 179.51	326.04 $\pm$ 126.57	47	0.05
$\Delta$ Duration (ms)	25.68 $\pm$ 16.79	30.96 $\pm$ 11.21	60	0.17

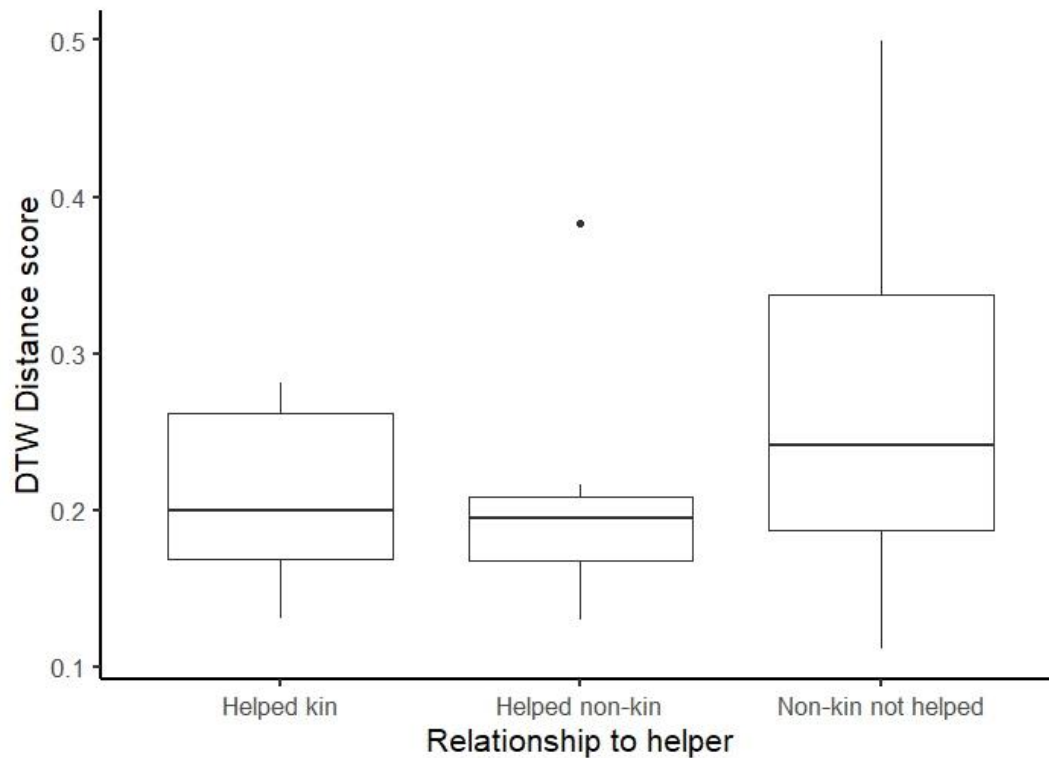


693 **Fig. 1.** Dissimilarity of the long-tailed tit churr call ( $n = 907$  calls from 54 birds) within  
694 and between individuals, measured using distance scores generated by dynamic-time  
695 warping analysis in *Luscinia*.





696 **Fig. 2.** The relationship between churr call dissimilarity and relatedness in long-tailed tits  
 697 based on dynamic time-warping analysis (DTW score) and the difference ( $\Delta$ ) in a range  
 698 of acoustic parameters. Mantel R correlations are shown for dyadic comparisons among  
 699 breeders based on (a) degree of kinship calculated from the social pedigree ( $n = 80$ ), and  
 700 (b) genetic relatedness estimates ( $n = 88$ ).



701 **Fig. 3.** Dissimilarity of churr calls among helper-breeder dyads: (i) helpers and related ( $r$   
702  $\geq 0.25$ ) breeding males they helped ( $n = 9$ ); (ii) helpers and unrelated breeding males they  
703 helped ( $n = 8$ ) and (iii) helpers and unrelated breeding males within helping range ( $\leq$   
704 750m) that they did not help ( $n = 218$ ). Call dissimilarity was measured using dynamic  
705 time warping analysis (DTW). Boxplots represent median  $\pm$  IQR. A full model output is  
706 reported in Supplementary Table S4.