



Peer Community In Ecology

Two common European songbirds elicit different community responses with their mobbing calls

Tim Parker  based on peer reviews by 2 anonymous reviewers

Ambre Salis, Jean Paul Lena, Thierry Lengagne (2023) Acoustic cues and season affect mobbing responses in a bird community. bioRxiv, ver. 5, peer-reviewed and recommended by Peer Community in Ecology. <https://doi.org/10.1101/2022.05.05.490715>

Submitted: 06 May 2022, Recommended: 28 February 2023

Cite this recommendation as:

Parker, T. (2023) Two common European songbirds elicit different community responses with their mobbing calls. *Peer Community in Ecology*, 100420. <https://doi.org/10.24072/pci.ecology.100420>

Published: 28 February 2023

Copyright: This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Many bird species participate in mobbing in which individuals approach a predator while producing conspicuous vocalizations (Magrath et al. 2014). Mobbing is interesting to behavioral ecologists because of the complex array of costs of benefits. Costs range from the obvious risk of approaching a predator while drawing that predator's attention to the more mundane opportunity costs of taking time away from other activities, such as foraging. Benefits may involve driving the predator to leave, teaching relatives to recognize predators, signaling quality to conspecifics, or others. An added layer of complexity in this system comes from the inter-specific interactions that often occur among different mobbing species (Magrath et al. 2014).

This study by Salis et al. (2023) explored the responses of a local bird community to mobbing calls produced by individuals of two common mobbing species in European forests, coal tits, and crested tits. Not only did they compare responses to these two different species, they assessed the impact of the number of mobbing individuals on the stimulus recordings, and they did so at two very different times of the year with different social contexts for the birds involved, winter (non-breeding) and spring (breeding). The experiment was well-designed and highly powered, and the authors tested and confirmed an important assumption of their design, and thus the results are convincing. It is clear that members of the local bird community responded differently to the two different species, and this result raises interesting questions about why these species differed in their tendency to attract additional mobbers. For instance, are species that recruit more co-mobbers more effective at recruiting because they are more reliable in their mobbing behavior (Magrath et al. 2014), more likely to reciprocate (Krams and Krama, 2002), or for some other reason? Hopefully this system, now of proven utility thanks to the current study, will be useful for following up on hypotheses such as these. Other convincing results, such as the higher rate of mobbing response in winter than in spring, also merit following up with further work.

Finally, their observation that playback of vocalizations of multiple individuals often elicited a more mobbing response than the playback of vocalizations of a single individual are interesting and consistent with other recent work indicating that groups of mobbers recruit more additional mobbers than do single mobbers (Dutour et al. 2021). However, as acknowledged in the manuscript, the design of the current study did not allow a distinction between the effect of multiple individuals signaling versus an effect of a stronger stimulus. Thus, this last result leaves the question of the effect of mobbing group size in these species open to further study.

References:

Dutour M, Kalb N, Salis A, Randler C (2021) Number of callers may affect the response to conspecific mobbing calls in great tits (*Parus major*). *Behavioral Ecology and Sociobiology*, 75, 29.

<https://doi.org/10.1007/s00265-021-02969-7>

Krams I, Krama T (2002) Interspecific reciprocity explains mobbing behaviour of the breeding chaffinches, *Fringilla coelebs*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269, 2345–2350. <https://doi.org/10.1098/rspb.2002.2155>

Magrath RD, Haff TM, Fallow PM, Radford AN (2015) Eavesdropping on heterospecific alarm calls: from mechanisms to consequences. *Biological Reviews*, 90, 560–586.

<https://doi.org/10.1111/brv.12122>

Salis A, Lena JP, Lengagne T (2023) Acoustic cues and season affect mobbing responses in a bird community. *bioRxiv*, 2022.05.05.490715, ver. 5 peer-reviewed and recommended by Peer Community in Ecology. <https://doi.org/10.1101/2022.05.05.490715>

Reviews

Evaluation round #4

DOI or URL of the preprint: <https://www.biorxiv.org/content/10.1101/2022.05.05.490715v4>

Version of the preprint: 4

Authors' reply, 27 February 2023

Dear recommender,

Please find the manuscript with a formatting helping scientific readers. We changed the word '15' to 'fifteen'. Thank you again for all the advice given to improve this manuscript.

Decision by Tim Parker , posted 21 February 2023, validated 23 February 2023

Minor revision requested

Dear Ambre Salis,

Thank you for your revisions. As soon as you post your re-submission, I will recommend this preprint.

I have one small editorial request - on line 282, please change "15" to "Fifteen" because it is the start of a sentence.

Sincerely,

Tim Parker

Evaluation round #3

DOI or URL of the preprint: <https://doi.org/10.1101/2022.05.05.490715>

Version of the preprint: 3

Authors' reply, 21 February 2023

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Tim Parker](#) , posted 07 February 2023, validated 09 February 2023

Minor revisions requested

Dear Ambre Salis,

Thank you for your revision. I am prepared to recommend this pre-print after you make a few small edits to improve clarity.

I look forward to seeing the final manuscript.

My suggestions follow by line number:

103: "playbacks" should be "playback"

also, "is" should be "in"

113: insert ")" after "hops"

250: This statement "The two main species, apart from coal and crested tits, were..." is, to me, misleading, since it suggests that both coal and crested tits responded more often than goldcrests, but from fig 1A, it seems that only crested tits were more common responders than goldcrests.

251: Here and elsewhere, I encourage consistency with bird name capitalizations. Given that you do not capitalize coal tit or crested tit, I suggest you not capitalize goldcrest, marsh tit, or others.

283: could you state the rates of occurrence at playback of coal tits and crested tits relative to goldcrests and chaffinches.

Also, I am confused – given the number of crested tits responding was higher than the number of goldcrests (according to Fig 1B), how is it that goldcrests responded 24% of the time and fewer than 25% of trials had any response? This suggests that goldcrests responded in (almost) every case where there was any response from any species – is that correct? Also, does this mean that most of the trials with crested tits responding had multiple individuals responding, and so the percent of trials with crested tits responding was not higher than the percent of trials with goldcrests, even though the total number of trials with goldcrests responding was higher than (or as high as) the number with crested tits?

319: change "per treatment" to "across treatments"

334: change to "middle graphs are responses"

380: change "modulates" to "modulate"

Also, I don't think we have sampled widely enough across birds to say "usually" here. Maybe say "often". Or maybe say something like "Where it has been studied, birds usually ..."

427: insert "a" before "community"

443: change "explore" to "explored"

483: insert "support for" so that this reads "we found support for different models"

Sincerely,

Tim Parker

Evaluation round #2

DOI or URL of the preprint: <https://doi.org/10.1101/2022.05.05.490715>

Version of the preprint: 2

Authors' reply, 05 February 2023

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Tim Parker](#) , posted 05 December 2022, validated 05 December 2022

Acoustic cues and season: request for additional revision

Dear Ambre Salis, Thank you for the revision of your manuscript "Acoustic cues and season affect mobbing responses in a bird community". I have examined the revision as has an independent reviewer (Reviewer #2 from round 1), and we are both pleased with your revisions, but we both also have some concerns that need to be addressed. I have three primary concerns. 1. My first concern relates to transparency of your reporting. You are reporting results only from your top models, and this practice leads to bias in the literature. Switching from frequentist to Bayesian model selection does not remove the need to report the outcomes of all models. Bias can emerge from selective reporting of top models selected through Bayesian methods as it can from frequentist model selection. Please include two additional tables (possibly in a supplement): A. A table with the model rankings and BIC etc for the full set of models you examined B. A table that includes the parameter estimates (and SE etc.) for all models tested 2. My second concern involves your interpretation of your results. You make several statements about the interpretation of your results that I think many readers will feel are too confident. For example: Abstract: "Our results therefore support the hypothesis that birds consider both the species and the number of callers when joining a mobbing chorus in winter" Although it is plausible that your method induced a differential response due to the perception that number of callers varied, with the current design, it seems we cannot rule out an explanation of duty cycle alone. However, I think you can build a more persuasive case that your results are not due to duty cycle alone. I write more about this below. Discussion: "These results corroborate the hypothesis that a greater number of birds mobbing a predator represents a lower risk for a potential mobber" I also think this claim is too strong (because there was no experimental manipulation of risk). Let's first consider the duty cycle argument. You make the case that overlapping of calls in playback reduces the chance that calls appear to be a single bird, which is plausible. However, that remains an untested assumption, and the alternative, that birds simply respond more often and more strongly when the signal is stronger is also plausible based on basic ethological theory. The duty cycle explanation is also a more parsimonious explanation. Of course, the duty cycle argument is not inconsistent with a role for selection in favoring responses to larger groups of mobbers, but assessing a role for responding based on group size would require different evidence. Ideally this evidence would be experimental and would control for duty cycle. However, because results differed between species between seasons, you could argue that birds are not simply responding to the signal with the stronger duty cycle - they do so in some cases, but in other cases, they do not. So, I think you should (a) make this case explicitly, (b) make sure you acknowledge, where relevant in the abstract and discussion, that duty cycle was a confound that limits the strength of your inference, and (c) suggest particular experimental designs (in the discussion) that would control for the duty cycle confound. Now on to the risk argument. If we believe the argument I just put forward in the prior paragraph, that the response differences you observed between seasons/species suggest that mobbers are doing something more than just responding differently in response to different signal intensity (duty cycle), we still don't have evidence regarding a role for risk in this behavior. It would be acceptable to present differential risk as a potential explanation for the outcomes you observed, but I would do so with more caution than you currently present. I would also suggest that you consider what evidence might be informative regarding the

risk hypothesis. For instance, maybe you could discuss an experiment that used mounts of different predator types with different threat levels or in locations in which they posed more or less of a threat, or provided different amounts of cover to the mobbers, or possibly some other manipulation that would allow you to assess if varying the threat level varies the mobbing behavior. Regardless, please be more cautious overall regarding your discussion of the applicability of various causal hypotheses.

3. My third concern is that one of your primary inferences involves a comparison between responses in winter and spring, but you do not actually test for this difference. You make the case that the responses in winter and spring are not comparable, but it is not clear to me that this is the case. If the overall response rates differ between seasons, then you can just fit different intercepts or slopes for different seasons in your model. At this point it seems that you have several options. The simplest would be to explicitly acknowledge, in the abstract and in the discussion, that you did not compare winter to spring responses statistically, and that you are making only a qualitative assessment that the patterns differ. Another option could be to conduct a post-hoc test based on a single model that included both winter and summer data to quantify the effect of season (and especially interactions with season) on the mobbing response rates of interest. The most thorough option would be a model selection procedure like the one you already conducted, but including both winter and spring data and season as a predictor variable in various forms (including presence and absence of interactions with season). I would find any of these three options reasonable.

I also want to call attention to several of the suggestions made by the reviewer in this round of reviews.

1. The reviewer had concerns regarding your implementation and reporting of hurdle models. Some of these concerns can be addressed by more thorough reporting of your methods. You should consider and respond to all of these comments.

2. It is my assessment that your reporting in Table 1 is correct (you have not reversed the occurrence and intensity results as suggested by the reviewer), but please check and be certain.

3. The reviewer asks for justification for why you include different terms in different components of the model (zero inflation vs. intensity) – this problem may be partly addressed by more thorough model reporting (as I mention above), but also suggests that you should devote more space to justifying the set of models you assessed.

Below are my specific suggestions for edits by line number.

36-39: please acknowledge the duty cycle difference between stimuli somewhere around here in the abstract so that the reader understands that number of callers was confounded by number of calls in the stimulus recordings.

41: Change “context interacts can strongly affect” to “context can affect”

95: I suggest you change “stability of acoustic cues” to “stability of response to acoustic cues”

96: I suggest you change “as much as” to “as well as”

108-109: This sentence “Each spot was selected close to a tree allowing birds’ approach and concealment of experimenters, following existing trails.” is somewhat confusing, and would be clearer as something like: “Each spot was selected along an existing trail but close to a tree allowing birds’ approach and concealment of experimenters.”

128-129: some clarification would be valuable here. Were the four tests at each spot carried out on the same day spaced 5 minutes apart, or was each consecutive test at a different spot, and each of the four tests at a spot within a season conducted on a different date? I think it was the second, but please clarify.

166: “NW-A45 Sony” - is this the speaker or ??

175: Thank you for adding this clarification: “The two observers agreed on the lowest number of birds seen simultaneously by both experimenters.” However, I think the wording needs revision. Wouldn’t this be the “highest number of birds seen simultaneously by both observers”? If the observers each saw 1 bird at the same time, and then a few seconds later, each saw 2 birds simultaneously, wouldn’t you count this a ‘2’ (highest) rather than ‘1’ (lowest)?

180: This statement: “Since the number of responding birds during the winter cannot be strictly compared to the one observed during the spring” is not informative, and is not obviously true. If you choose to keep the spring and winter analyses separate (see my comments about this above), I suggest you change the wording here to something like: “Since social conditions for our study species differ between winter and spring and factors influencing rates of response presumably therefor differ ...”

182: I suggest a change from “at the community level...” to something like “of any species (“community level”)...”

186: insert “us” before “to take”

186: insert “an” before “excess”

187: change to “zeros”

188: change “a first” to “an initial”

188: change “determine” to “determines”

190: change “determine” to “determines”

195-198: I agree with the reviewer

that more information is needed here 198: I found this confusing. To help clarify, I suggest you change “the one of the number of callers and...” to “the effect of the number of callers, and ...” [note word change and the addition of the comma] 210: change “calculate how much better is the best model compared to the other ones” to “calculate how much better the best model is compared to the other ones” 212: although I agree that delta AIC or delta BIC is often used a threshold, it is rarely referred to as a ‘significance’ threshold, so you may not wish to use this word. Instead you might say something like “models with a delta >2 are commonly considered to have substantial support ...” 293: some additional information would be useful in these figure headings. Close inspection and consideration leads me to conclude that categories are stacked in each bar (rather than layered, so for instance, the number of crested tits in spring responding to the 1CO treatment is about 25 rather than slightly more than 50. If I am correct, then wording could be added as a new sentence on line 295 to say something like “Responses to each of the four treatments are stacked in sequence on each bar so that the entire bar represents the sum of all responses by a given species per treatment” 324: somewhere (maybe in a supplement), you should include the full set of models examined and their model ranking statistics (BIC etc.) AND ALSO their associated parameter estimates and corresponding standard errors. 326: clarify what ‘further reduced’ means (your method of model selection) 327: change “the one of the number of callers as well as” to “the effect of the number of callers, as well as” 353: I think the word “corroborate” here is not ideal. I would prefer something like “These results are consistent with the hypothesis...” I prefer this wording because your results are equally consistent with other hypotheses. For instance: (a) a stronger signal is more likely to be detected by potential additional mobbers and is therefore more likely to attract more mobbers (b) a stronger signal is more likely to reach the threshold necessary to trigger mobbing in an individual. I will point out that both (or either) of these two hypotheses (a and b) could be true while your hypothesis is true, but both a and b could be true while your hypothesis is not true. I think it would be useful to discuss what additional evidence you would want to examine to evaluate the plausibility of your relative risk hypothesis. 398: I suggest changing “opposition” to “contrast” 402: I suggest changing “aggressivity” to “aggressiveness” (here and elsewhere in the paper) 406: This would be more clear to the reader if you again explained how “occurrence” differed from “intensity” here (like you do in the figure 2 and 3 headings). Otherwise, it is not immediately obvious to the reader how what you have written here differs from what you wrote on line 400. One way to do this would be to write something like “Additionally, not only did fewer individuals respond in spring than in winter, but in spring, the proportion of locations with any response was lower than in winter” 407: I think you mean “populations” (plural) here since you are taking about the populations of multiple species 441: change “despite our” to “despite the fact that our” 449: change “in adequacy with” to “consistent with” or “similar to” 455: add comma after “We have” and after “however” 457: change “are” to “is” (because “status” is singular) 459: delete “very” (it is not necessary to make your point)

Reviewed by anonymous reviewer 2, 21 November 2022

21 November 2022

I thank the authors; they have done great work addressing all the concerns and queries I raised in the first round. The statistical method used in conjunction with the experimental design (factorial design) is now appropriate. The preprint organisation has also been improved to a level the reader could easily follow.

I have made a couple of suggestions to improve the presentation of results and have a query of the models in Table 1. Otherwise, the preprint is scientifically sound and may require editing before publishing in PCI or any other journal.

Major comments:

In Table 1, mobbing responses are listed under two different response variables: response occurrence and intensity. As far as I understood, and what I see from model syntax in the R script, the zi- zero inflation part reported under the mobbing intensity.

In fact, the zi part of the model helps to define what variables contribute to zero inflation; it could be one variable than the other, or both variables equally contribute to zeros inflation. As the authors mentioned, there is no distinction between true and false zeros in Hurdle models. However, for example, the authors need to justify why they think that, Emitter species + number of callers contribute to zero inflation in one model and, why only the Emitter species contribute to zero inflation in another model (this is an example, it may apply to all the models presented in the table).

In addition, authors may consider defining theta in the table caption or the statistical methods section where appropriate; otherwise, people who think in a Bayesian way might be confused with parameter estimates.

There are a couple of issues that need to be resolved or need explanation here:

01. I agree that the presence of excess zeros does not warrant using zero inflation models. Hurdle models can use alternatively, but clarification is needed on how the zero (inflation) arises in three different analyses. The first step is to evaluate the overdispersion (either with Poisson or Negative binomial distribution), which may be done using a simulation test. If the authors did overdispersion tests before selecting the Hurdle model procedure, please mention it on Page 9, lines 185-188.

02. In this study, as far as I understood, zeros (inflation) may arise on two processes: 1 zero may occur when the species do not present at the point or species present at the point but did not show any mobbing responses. For example, in community-level analyses, the number of responding birds may represent single species or different species (i.e., if four birds showed a mobbing response, it could be from a single species or four different species).

Were zero responses specifically generated, either absence of the mobbing species or no mobbing response towards the soundtracks, particularly for coal tits and crested tits? Clear identification of the zero-generation process is also essential when defining the glmmTMB models.

03. Using the same terms in both factorial design and the models makes sense. Model selection may help choose different distribution fittings (i.e., negative binomial vs Poisson) while keeping the fixed effects in the model. This could also extend to test the hypothesis of the additive vs interaction effect of the same model instead of dropping or adding terms.

04. Why is glmmTMB control "BFGS" used in models? I presume this is because to account lack of convergence in some models. If that is the case, please include the relevant details which would be helpful for the reader in the statistical methods section.

Zuur,A.F and Ieno,E.. Beginner's Guide to Zero-Inflated Models with R. 2016. (Chapter 6 for Hurdle models). This book extensively discusses the statistical and practical background and updated version of the methods introduced in Zurr et al. 2009, which authors cited.

Minor comments:

Page 4, line 86: rephrase or add (unclear)

Page 4, line 88: If the authors can back this with a reference, that would be great.

Page 5, line 110: what is 'X'

Page 6, line 112: please give the breakdown (n=22, coal tits? great tits?).

Page 9, line 179: R version 3.6.1 was released in 2019, not 2022. Please ensure all the package versions used in the preprint are correct and include their references.

Page 9, lines 195-197: Is the overdispersion the main reason to use negative binomial distribution??

Page 10, line 208: if it is due to sampling size, then AIC corrected is also helpful; perhaps it may be unobserved heterogeneity. Brewer et al. 2016 Methods. Ecol.Evol. Volume 7(6) p.679-692.

Page 10, line 200: were random effects introduced as an intercept?

Page 11, line 225: it would be helpful to provide contrast measures using the final model. The authors may use the emmeans package (Ver 1.8.2), 2022 or an equivalent package to get contrast estimates. Estimates also provide strong statistical evidence for the graphical presentations in Fig1-3.

I am sure the following reference: Ratnayake et al. 2021. Behav.Ecol Vol 32(5) pages 941-951. It may be helpful to include some necessary information in the methods section, and please note that this is not an indication to cite the reference. However, the reference may be relevant as the study used mobbing calls of noisy minors to test the occurrence and the intensity of the responses of Australian magpies.

I hope comments will help improve the preprint quality, particularly parts in the statistical methods section. Finally, I congratulate the authors for their good work.

Evaluation round #1

DOI or URL of the preprint: <https://doi.org/10.1101/2022.05.05.490715>

Version of the preprint: 1

Authors' reply, 04 November 2022

[Download author's reply](#)

[Download tracked changes file](#)

Decision by [Tim Parker](#) , posted 03 August 2022

Acoustic cues and season: request for revision

Dear Ambre Salis,

I apologize for the delay in providing these comments. By the time I received the second review, I had left on holiday for much of July.

Two independent reviewers and I have read your manuscript (Acoustic cues and season affect mobbing responses in a bird community). The reviewers and I all see value in this study. I appreciate the use of a thorough design to simultaneously evaluate multiple potential influences on the response to mobbing calls, the large sample of trials, and the evaluation of an important assumption of your experiment. However, we also all identified some important areas where improvements are merited. I have provided some detailed comments below and both reviewers provided detailed suggestions as well. Please carefully consider all these suggestions and either implement the suggestions or explain why you have not done so if you chose to resubmit a revised manuscript.

Before providing my detailed comments and those of the reviewers, I want to call attention to several points that are particularly important.

Both reviewers noted that group size may not be expected to correlate with the reliability of mobbing calls. I encourage you to explore the literature on this subject further, and to update your discussion of this topic.

Reviewer 1 and I both felt that your explanation in the main text of the supplementary experiment (done to assess the likelihood of overlapping responses between playback locations) is insufficient. I encourage you to either bring this experiment to the main document, or at least to provide more details in the main document.

Reviewer 1 and I also wanted to see more information about the 3-bird playback stimuli. Besides addressing the questions of Reviewer 1, I encourage you to consider adding the stimuli and the sound spectrograms of the stimuli to your supplement.

On a related topic, both reviewers and I share the concern about your interpretation of the 3-bird stimulus due differences in duty cycle between the treatments.

Reviewer 2 identified several important issues related to your statistical analyses, including a model convergence error (which I also replicated when I ran your code). I do not believe there is only a single correct way to analyze a dataset, but I would like you to seriously consider Reviewer 2's recommendations and concerns.

Sincerely,
Tim Parker

What follows are some specific concerns that I noted as I read the manuscript (organized by line number):

41: "We therefore confirm the hypothesis" – should be something like "We therefore find support for the hypothesis"

161-162: when a bird was detected in the area, what was your protocol? Did you wait for it to leave or move to another location?

113: How did you determine the order in which you visited survey locations?

159: How did you resolve differences in observations between observers?

174: please explicitly state that you excluded cases of zero detections from the intensity analyses

177: which versions of the lme4 R package? Also, please cite the package (e.g., 'Bates et al....' for lme4).

179: if you decide to remain with analyses using Poisson error (instead of taking the suggestions of reviewer 2), you should report the details of your test for overdispersion.

183-185: This sort of step-wise procedure can lead to biased model estimates (see Forstmeier and Schielzeth. 2011. Behav Ecol Sociobiol 65:47-55). Given that you present full models, it is not clear to me why you were using a step-wise procedure.

189: again, more information needed about packages

Table 1 vs. Table 2. The first table in the text is labeled 'Table 2', but presumably it should be labeled 'Table 1' – it seems that when you cite Table 1 in the text, you are referring to the table currently labeled 'Table 2'

Table 2 (as currently labeled): Why do you not present number of mobbers results for coal and crested tits? You seem to have that information in another form in Figure 2.

Table 1 (as currently labeled): The heading says "propensity and intensity", but the content of the table looks like only one or the other, but definitely not both.

Table 1 and Table 2/ Figure 2: You currently only present the estimates in graphical form. It would be useful to present the actual numbers (either in an expanded form of Tables 1 and 2, or maybe in a table in the supplement).

Table 2: 'NB' should be defined in the heading

338: You should more explicitly acknowledge the limitations of your experiment here (duty cycle not standardized).

Reviewed by anonymous reviewer 1, 23 May 2022

[Download the review](#)

Reviewed by anonymous reviewer 2, 10 July 2022

[Download the review](#)