

This interesting experimental study tested whether Great and Blue Tits respond to heterospecific alarm calls of sympatric and allopatric titmice species. Using playback at a feeder the author showed that both study species are able to recognize and respond to at least some heterospecific alarm calls.

The text is generally well written and reads nicely. The Introduction is appropriately structured and introduces into the topic very well. The results are properly discussed. However, there could be a better fluency and connection among sentences in some parts of the text. I am not an expert in this field of research but I suspect that some references were cited improperly (see two examples in comments below). There also remain few concerns regarding methods and results, which need to be carefully addressed by the author. I am confused by a statement in the Conclusions section that both studied species did not respond to any heterospecific alarm calls despite results and conclusions itself clearly show otherwise.

It is an incredible feat to conduct all the experiments on March 2021 and defend the thesis two months later, hats off! The results show interesting patterns but, in my opinion, more experiments are needed before the results can be published in a scientific journal. I also raised other methodological and statistical points, which need to be carefully reviewed before considering this study to be submitted as a manuscript. The supplements are very nice and useful. I think that there are few appropriate spots in the Discussion, where references to supplementary figures could be also included.

In conclusion, I like the study and its original idea and after correcting few weaker points I believe it can be published in a good scientific journal. I assume that far-reaching negative effects caused by the coronavirus pandemic might affect directly or indirectly this study (as did my research) and, thus, I propose the rating “very good” (velmi dobře).

Page 8, Methods – subsection Study site: I would like to see a more detailed description of the feeder being used, including its dimensions. Also, more information on camera set up is needed (model type, positioning, concealment).

It is unclear to me how many experimental bouts (1 exp. bout = 75 experiments/playbacks) were performed within the two weeks of data collection. The methods and results rather suggest that data collection was performed within a very narrow time window (maybe only one day?) and I think that a careful explanation is necessary to justify such approach. For example, on page 14, 2<sup>nd</sup> paragraph, last sentence you admit that “As we ran our experiments in March, this can be another reason that could affected our results.”. Specifically, you argue that Blue Tits and Great Tits may be less interested in heterospecific alarm calls in the late winter and, thus, experiments throughout the whole winter period, or at least periods other than late winter might be preferable.

Page 8, Methods – subsection Playbacks: Xeno-canto recordings of the same call type in the same species often considerably vary in many characteristics but it seems that no effort has been done to standardize the quality of recordings. Why? How long did the *song* recording last? What software has been used for editing?

Page 8, Methods – subsection Experiment: What was the position of the experimenter? How far was she from the feeder? Did she use a hide?

Page 9, Methods – subsection Recorded data and analyses: I am not sure if I understand well to the phrase “one minute before and one minute after the stimulus played”. Does this really mean that birds were counted exactly one minute before and after the playback? I think it rather means the total number of birds at feeder *during* one minute before/after the playback? As the birds were not marked, does not make more sense to count the *maximum* number of Blue and Great Tits present at a time within the minute before and after the playback? I think that this approach would better account for pseudoreplication issue and a possibility of some individuals being more active than others (like re-visiting the feeder more or less often due to the personality, hunger level, ...).

I am not sure if choosing a linear model to test proportions is a correct approach. The main reason is that a linear model, regardless of the transformation being used, allows to make model predictions  $<0$  or  $>1$ , which does not make sense for proportion data (restricted to the interval 0 and 1). Why was not employed a more appropriate statistical model for this analysis, e.g., a generalized linear model using some of binomial link family?

Pages 10 – 12, Results: There is no information about sample size in the text or figures. In general, it remains unclear to the reader how many experiments and how many times were performed for each treatment in total in this study.

Page 13, Discussion, 3<sup>rd</sup> paragraph: The finding of a weak response in Great Tits to the Blue Tit alarm call is interesting. It would be interesting to see also reactions of Blue Tits to the Great Tit song and alarm calls. I think that

such results would be significantly beneficial for the study. Why these Great Tit song and alarm calls were not considered at all?

Is the citation of Dutour et al. (2021) appropriate here? I think that responses of Great Tits to Blue Tits and Chaffinches has been studied elsewhere (e.g., Randler and Vollmer *Naturwissenschaften* 2013) but not by Dutour et al. (2021).

It is indeed surprising that Great Tits responded weakly to the Blue Tit alarm call. As I am not clear about sample size of this study, I would like to know if the power of this analysis was satisfactory or there is a need for collection more data.

Page 14, 1<sup>st</sup> paragraph: I am not sure that reference of Dutour and Randler (2021) is cited in the right context here. Please, re-check this because their Abstract states “great tits responded similarly to mobbing calls from several heterospecific callers as they did toward mobbing calls from one heterospecific caller”.

Page 16, Conclusions: I read the sentence “We may conclude that Blue Tits and Great Tits do not respond to any alarm of any titmice species.”. I do not understand this sentence because Results clearly show that both Blue and Great Tits significantly responded to at least some of heterospecific alarm calls. Also, in the Conclusions section itself I can read that “Both species strongly responded to American Black-capped Chickadee” or “Great Tits also responded to alarms of Varied Tits”.

Abstract is missing. This is not a mandatory part of the master thesis?

#### **Minor comments:**

Page III and Table of contents: misspelled word “AKNOWLEDGEMENTS”

Page 1, Introduction – subsection Alarm calls, 3<sup>rd</sup> paragraph: Unusual and imprecise biological phrase “selfish selection pressure”. I see that this term was used by Hollen and Radford in *Anim Behav* 2009 but this is also probably the single published paper which used this improper phrase.

Page 2, Introduction – subsection Alarm calls, 4<sup>th</sup> paragraph: wrong phrase “it can increases”

Page 3, Introduction – subsection Information coded in alarm calls, 2<sup>nd</sup> paragraph: Incorrect phrasing “...and if it decreases in frequency represents an increase in hostility.”

Page 4, Introduction – subsection Heterospecific alarm calls, 1<sup>st</sup> paragraph: misspelled word “scape”

Page 6, hypotheses: in the first hypotheses correct wording to “only to”

Page 7: correct “...primarily focused on...” to “primarily focus on”

Page 9, Methods – subsection Recorded data and analyses: Does t-test refer specifically to Welch’s t-test, which considers unequal variances? What does “R 92” mean when citing R software?

Page 13, Discussion, 1<sup>st</sup> paragraph: Maybe better to delete this paragraph and use these sentences elsewhere in the text.

Page 14, 2<sup>nd</sup> paragraph, last sentence: wrong phrase “could affected”

References: probably some typos in the full reference of Ekman (1989) and other minor inconsistencies (Latin names should be in italics, some author names are in capital letters only, ...) to be corrected before manuscript submission.

Ve Znojmě, 17.5.2021  
Peter Samaš