

Learned use of an innate sound-meaning association in birds

Corresponding Author: Dr William Feeney

Version 0:

Decision Letter:

18th November 2024

*Please ensure you delete the link to your author homepage in this e-mail if you wish to forward it to your co-authors.

Dear Dr Feeney,

Your Article, "Learned use of an innate sound-meaning association in birds" has now been seen by five reviewers--given the broad methodological toolkit used in the paper, this number of reviewers was necessary to make sure we had all components assessed. You will see from the referee comments below that while they find your work of considerable potential interest, they have raised quite substantial concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication, but would be very interested in considering a revised version that addresses these serious concerns.

We hope you will find the reviewers' comments useful as you decide how to proceed. If you wish to submit a substantially revised manuscript, please bear in mind that we will be reluctant to approach the reviewers again in the absence of major revisions.

In particular, no reviewer yet seems to be convinced that the calls are functionally referential, however reviewer 2 has some suggestions of additional qualitative data/experiments that might help strengthen this interpretation. We recognise that these are likely to represent a substantial undertaking, but will be necessary to improve the paper's robustness. If you do wish to revise your paper and resubmit to Nature Ecology & Evolution I would also recommend that you follow reviewer 5's advice on restructuring so that the findings rest on their own merits, rather than within a historical framework.

If you choose to revise your manuscript taking into account all reviewer and editor comments, please highlight all changes in the manuscript text file [OPTIONAL: in Microsoft Word format].

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

If revising your manuscript:

* Include a "Response to reviewers" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

* If you have not done so already we suggest that you begin to revise your manuscript so that it conforms to our Article format instructions at <http://www.nature.com/natecolevol/info/final-submission>. Refer also to any guidelines provided in this letter.

* Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

Please use the link below to submit a revised paper:

Link Redacted

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. If you cannot send it within this time, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Ecology & Evolution or published elsewhere.

Nature Ecology & Evolution is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. This applies to primary research papers only. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit www.springernature.com/orcid.

Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further.

Thank you for the opportunity to review your work.

Yours sincerely,

[redacted]

Reviewer expertise:

Reviewer #1: bird macroevolution and phylogenetics

Reviewer #2: animal communication, vocal learning, playback experiments

Reviewer #3: vocal learning in birds, playback experiments

Reviewer #4: avian brood parasite ecology

Reviewer #5: avian brood parasite ecology

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

This manuscript compiles experimental evidence to support the claim that vocal responses by host species to their brood parasites are convergent across large sections of the bird phylogeny. The authors argue that this is of interest because it shows learned use of a referential vocalisation coupled with an innate response, providing some support for one proposed mechanism for the evolutionary origin of symbolic associations in human language. Overall, this work deals with an interesting topic and can be commended for many strengths, including excellent figures and high-quality writing. Nonetheless, I ended up in two minds about the suitability of the paper for this journal because of a number of weaknesses in the content, study design and interpretation.

The evidence brought together by the authors comes across as slightly haphazard, with a number of small experiments, conducted in widely different regions, designed to fend-off specific critiques. While this strengthens the manuscript to a degree it comes across as slightly confected – an amalgamation of multiple supportive yet slightly inclusive lines of evidence. The number of twists and turns in the storyline also gives the sense of a rather specialised paper with lots of methodological detail required to carry the narrative.

My main concern in terms of the quality of evidence supporting the main claims was as follows. The whining call, in my experience, is not limited to the context of brood parasites. It seems to be used more generally as a warning signal in the context of risks to the nest contents (rather than to adults). It may be elicited by predators such as snakes or rats that pose minimal risk to the parent birds, or, in some bird species, even by people closely approaching nest sites. I would question whether these vocalisations are referential to brood parasites alone. The manuscript does attempt to fend off this type of criticism with an experimental test involving presentation of a supposed nest-predator (snake) near active nests. While this treatment did not seem to elicit whining notes, the sample sizes are not great and I am not sure that an immobile snake model elicits the same response as the more slithering variety.

If the whining call type is a more general signal of risk to nests or nest-contents then it seems a bit harder to argue for direct symbolic association between the vocal signal and the context of brood parasitism. Both nest-predators and brood parasites can be mobbed by conspecifics (and heterospecifics), so a more general function of the whining call to communicate risks is possible. It is necessary to consider whether we, as humans, may be superimposing the meaning of 'brood-parasite' onto the sound, whereas the organism is using the call as a catch-all signal to convey a variety of risks to nest contents. Brood parasites are then potentially batched together with nest predators in the sound symbolising "risk" and species are able to

learn new risks, be they from parasites or predators. If the innate response to nest-risk signals is universal, then the process of learning which species pose risks – whether through predation or parasitism – can also be viewed as equally universal, rather than convergent as claimed in the paper. I think the text needs some editing to add more caution to the conclusions, and to acknowledge that the function of the whining signal requires further testing in a broader array of host species, and with more a convincing experimental set-up, including treatment with a larger sample of possible nest-predators.

Another concern is the assumptions around the category of non-hosts. Interpretation of statistical differences between current hosts and non-hosts is potentially problematic because the identity of host species undergoes turnover through time. Brood parasites move from host to host depending on arms-race dynamics. That is, when one host develops strong anti-parasite defences the brood-parasite benefits from switching to a naïve host. One upshot of this is that what may seem like a non-host may well have been a host in its recent evolutionary history, therefore retaining innate or learned responses to brood parasites and warning signals from conspecifics or heterospecifics (the ghost of parasitism past). In addition, the distinction between hosts and non-hosts may be blurred by biases in knowledge, with some hosts overlooked in data-poor cases. Some discussion of the potential pitfalls of inferring or assuming non-host status on the basis of current information is warranted.

For the reasons above, I think the authors need to dial down the level of certainty they ascribe to the findings and conclusions. Despite these shortcomings, the overall message of this manuscript remains potentially interesting. The analyses provide incomplete but suggestive evidence that signals eliciting innate responses can be used as referential symbols, and applied to learned objects. The suggestion that this type of scenario could provide early building blocks for the development of more complex language is intriguing, but highly speculative. Whether the finding is as novel as the authors state needs to be checked by experts in related fields, including linguistics.

Reviewer #2 (Remarks to the Author):

This study investigated whining vocalizations produced in response to brood parasites across multiple passerine species. The authors conducted playback experiments to assess whether whining vocalizations are functionally referential by comparing the number of whining calls produced in response to brood parasite vs. predator mounts and by comparing the behavioral responses to playback of whining calls vs. anti-predator alarm calls. Additional playbacks of conspecific and heterospecific whining calls found that birds responded equally strongly to all whining calls regardless of whether it was their own species' call or not and regardless of whether any local species produce whining calls. Passerines (which comprise most brood parasite hosts) responded more strongly to whining call playbacks than non-passerines. Finally, Purple-crowned Fairy-wrens produced fewer whining calls in response to cuckoo mounts in years with less brood parasitism, a similar finding to previous studies on Superb Fairy-wrens, which the authors interpret to mean that the context of whining call production is learned. The authors also conducted phylogenetic analyses to show that whining calls evolve several times slower than antipredator alarm calls, and that the interspecific cooperation has a stronger effect on whining call production than intraspecific cooperation. The authors concluded from these results that whining calls are a referential call whose production is learned but that elicit an innate response across diverse species. They suggest that this provides support for the hypothesis that semantic signals can evolve from innate associations between sound and meaning.

I commend the authors on an incredibly comprehensive study in which field experiments conducted at multiple sites around the world were combined with acoustic and phylogenetic analyses to address an important evolutionary question. The question of how sound-meaning associations is likely to be of great interest to scientists from a variety of disciplines, including behavioral ecologists, evolutionary biologists, and linguists. This study attempts to address this question in a way that few others have by quantitatively investigating the evolution of referential calls. The manuscript is very well-written and lays out a compelling rationale for the study as well as clearly explaining the experiments and analyses that were conducted. The experiments were well designed and included proper controls (although I am not sure if the response measured was the most appropriate for the research question- see below). The figures are extremely well done and informative. I especially like Figure 1. I cannot comment on the methodology for the phylogenetic and linkage density analyses as this is outside my area of expertise.

However, despite my overall positive opinion of the paper I do have a concern about the central claim that whining calls are functionally referential and learned. The authors argue that their findings are especially novel and important because there are few if any previous examples of functionally referential calls with learned production and an innate response, and finding an example of such calls would support the hypothesis that symbolic meaning arises from pre-existing innate associations. But based on the evidence provided in this paper, I am not completely convinced that whining calls are functionally referential or learned at all.

To be considered functionally referential, a call must refer to a specific external stimulus, rather than simply being a vocal manifestation of the caller's internal emotional state. To demonstrate that a call is functionally referential, it is necessary to show not only that the call is produced in a particular context, and not only that the call elicits an appropriate response, but also that the response is specific to the supposed referent of the call (i.e., the call elicits a qualitatively different response than other calls produced in other contexts).

In this study, the authors played back whining calls and anti-predator alarm calls and found that the whining call playbacks resulted in more individuals (both conspecific and heterospecific) approaching the speaker compared to the anti-predator alarm calls. One possible interpretation of these results is that the whining call is a functionally referential call for brood

parasites. But an alternative interpretation is that the whining call is simply a high-urgency alarm call, rather than being referentially specific to brood parasites. I don't think it is possible to distinguish between these possibilities with the data provided because the difference in response to whining calls and alarm calls was quantitative rather than qualitative. The birds performed the same behavior (approaching the speaker) in response to both playbacks; they just did it more in response to the whining call playbacks. If brood parasites are a higher-urgency threat than predators (which is consistent with the observation that birds mobbed brood parasite mounts more than predator mounts), this would explain why more whining calls were produced in response to brood parasite mounts. It is worth noting that whining calls were occasionally produced in response to the predator mounts too, which suggests that whining calls are not wholly specific to brood parasites.

To convincingly show that the whining call is functionally referential to brood parasites, I think the authors would need to provide some data showing that birds respond in a qualitatively distinct manner to whining calls and alarm calls. I can think of a few different ways to do this:

- 1) Show that whining calls elicit a unique behavioral response that is appropriate to brood parasites but not to other threats. For example, yellow warblers respond to anti-cowbird seet calls by rushing back to their nest to prevent the cowbird from laying an egg in it. This is qualitatively different from the response to a predator. If whining calls elicit a similar type of behavior, this could be evidence that whining calls are truly specific to brood parasites.
- 2) Conduct playbacks both during and outside of the breeding season. If whining calls are truly specific to brood parasites, we might expect that birds would show a decreased response to whining calls outside of the breeding season when brood parasitism is not an issue. By contrast, the response to anti-predator alarm calls should be constant in and out of the breeding season since predators are a threat year-round.
- 3) Do a prime-probe experiment similar to Zuberbühler et al. 1999. First "prime" the birds by presenting a mount of a brood parasite or a predator. Then once they have seen the mount, "probe" them by playing back whining calls or anti-predator alarm calls. If the birds mentally associate each call with a particular stimulus, then they should show a decreased response to whining calls after being primed with a brood parasite mount but not after being primed with a predator mount, and they should show a decreased response to alarm calls in the opposite scenario. But if both calls are just generic warning calls with the whining call being higher urgency, then they should show the same decrease in response to the call playback regardless of which mount they were primed with.

Another concern I have relates to the claim that birds must learn the context in which to produce whining calls. To support this claim, the authors cite two previous publications (Langmore et al. 2012; Feeney and Langmore 2013). They also present an experiment in the supplementary material in which Purple-crowned Fairy-wrens produced fewer whining vocalizations in response to cuckoo mounts in years when brood parasitism was low, suggesting that fairy-wrens who have more experience with cuckoos are more likely to produce whining vocalizations in response to them. However, as far as I can tell, the cited studies and the Purple-crowned Fairy-wren experiment only suggest that fairy-wrens must learn that cuckoos are a threat. They don't directly address the question of whether whining calls themselves are learned. If whining calls are a generic high-urgency call rather than referring specifically to cuckoos, then it is possible that fairy-wrens innately produce whining calls whenever they feel extremely threatened or upset, and the only thing they need to learn is that cuckoos are a high-level threat.

Other comments

- For your analysis of macroecological correlates of whining call production, how did you definitively determine that a species does NOT produce whining calls? Did you have comprehensive data on the vocal repertoires for all these species? Is it possible that some of them do produce whining calls but it simply hasn't been documented yet?
- Line 243: You say that "more individuals from other species responded", but it is not clear from the way this sentence is worded what they responded to. Which playback stimulus elicited a greater response by other (non-yellow warbler) species?
- Lines 335-336: You say that out of 24 species you had 17 with 2 recordings per context and 8 with 1 recording for one or both contexts. But $17 + 8 = 25$, not 24. What explains this discrepancy?
- Lines 358-359: What features were included in this LDA (the one with 85% accuracy)? Was it the first 10 principal components? Please clarify this in the text.
- Line 524: How do you know that 30 minutes was enough time to wait in between trials? It doesn't seem like much time

Reviewer #2 (Remarks on code availability):

When I tried to access the data and code through the link provided to Dryad, the link did not work. Also, I do not have the requisite expertise to assess the phylogenetic methods used.

Reviewer #3 (Remarks to the Author):

Overall assessment:

Using a variety of methods, the MS shows that whining calls produced by hosts of brood parasites are similar acoustically and functionally across species. The study combines network analysis with playbacks experiments and attempts to come up with a unified theory about these signals as functionally referential. I am not sure I agree with the argument about referential signaling and its relations to evolution of human spoken language, I think these claims are somewhat speculative, but I find

the results interesting nonetheless. Overall, the study is well conducted and well presented, but several issues in the statistical analysis worries me a bit and requires careful attention before this MS can be published.

Major concerns:

1. A major argument in the MS is that species that occur in areas where the underlying brood parasite–host interaction network shows low interactions per species, rarely produce whining vocalizations. But I thought that, by design, network interactions are computed, at least in part, based on vocal interactions, is this correct? If so, this entire claim feels a bit tautological. Authors should clarify how is host interaction network computed, and how they validate independence.
2. Certain claims seem unsupported by data presented. For example, the title of figure 2 is “Whining vocalizations are functionally referential in multiple species”, but in fact, all the figure shows is response to playbacks in two species (one of which has been studied earlier). The rest of the argument is based on panel K, which only show a list of neighboring species that interacted as well, but with no quantification. In fact, I could not understand much from panel K. Overall, I think that the MS can improve if authors should show some level of modesty in their claims.
3. There is some confusion in my mind about many of the statistical claims made in the MS. The authors have a bad habit of presenting p values and chi2 without presenting the means for each group. Furthermore, colormaps are often not explained and are not, in any direct manner related to raw measures that can be easily understood. See specific comments below.

Specific comments:

1. Acoustic analysis:

This statement confuses me:

“We analyzed the acoustic features that distinguish whining vocalizations from alarm vocalizations and found that the linear discriminant function loaded most strongly with PC3, which was negatively associated with frequency bandwidth and positively associated with spectral flatness.”

If this is a new measure, perhaps it would be better to develop it on a training set and testing it on a different experimental data set? Otherwise, it is sort of ad hoc, and statistical significance of this acoustic distance is not robust. Please explain.

This one through me off:

“Using the function `lda` MASS v7.3-6050, we generated a single linear discriminant function, which correctly assigned 85% of vocalizations to context”

what is this function exactly? It is not enough to say you used a certain toolbox, you need to document the exact procedure here

Same here:

“A species was considered to produce distinctive vocalizations in parasite and predator contexts when the linear discriminant scores of their vocalizations differed in post-hoc tests run using `emmeans` v1.8.752.”

Seriously, is this something a reader should be able to grasp? Here again, I don't understand the procedure, it feels as if the authors are not motivated to help the reader to understand what they did but expect reader to trust that the methods are okay. Most concerning is the binary decision made here: why binarizing the data rather than using an acoustic distance metric derived from PC3?

Finally, reading this statement:

“The linear discriminant function loaded most strongly with PC3 (Data S6), which, considered alone, correctly classified 91% of vocalizations”

This is, in fact, higher than your discriminant function! I wonder if the MS can improve if authors just use PC3 as a metric – as is – no categorization, etc. A color map of PC3 would be much easier to interpret in my view. As it is, in Fig 1 (The global distribution of distinct or non-distinct vocalizations towards brood parasite), I cannot understand this color map. What is exactly corresponding to these colors? I would find it much easier to interpret if these color map values are simply the mean PC3 values in each location.

Similarly, here:

“Of the 24 host species analyzed, 22 species produced distinct vocalizations towards brood parasites”

This is a highly derived measure. It is sufficient to refer to the PC3 values in figure 1 (much more convincing in my view).

Some points than need further clarification:

“If within-species cooperation is a primary benefit of whining vocalization production, species that produce them should be more socially cooperative^{26,37}. Alternatively, if between-species cooperation is a primary benefit of whining vocalization production, species that produce them should reside in areas that have complex brood parasite-host interaction networks”
Why cooperative breeding helps here exactly? I guess in general, cooperative breeding allows more guarding, but even just having a colony of close by nests could help here. So I feel that within and across species networks are not easily distinguishable here. Everyone gains from the warning whining, and the more proximal the nests, the more they gain, isn't it right?

“complexity of underlying brood parasite–host network”

Please explain the proposed mechanism here, complexity is a complex term...

“opportunity for between-species cooperation”

How is this quantified exactly?

“Fig 2E: Distribution of species that attended cuckoo model presentation experiments or were detected in the vicinity of host nests. Blue shading indicates whining”

There is no blue shading in K!

Finally:

The statistical tests should all show mean values for each group, not only chi sq. This should be consistently presented in the entire MS

Reviewer #4 (Remarks to the Author):

Review of "Learned use of an innate sound-meaning association in birds"

The authors present a series of experiments to test whether whining vocalisations associated with a brood parasite in one songbird system evoke a response to that call in other songbird systems in different geographical areas - even if the last common ancestor was 53 million years ago. The paper fills an important gap in our conceptual framework of the phylogenetic history of learned sound associations. The authors show that some innate vocalisations elicit a similar response in other (naive) bird species in allopatry. The authors conclude that their findings are parsimonious with the view that learned sound symbolism derives from innate origins.

The main weakness of the manuscript in its current form is the treatment of alarm calls. In line 106, the authors write “alarm vocalisations (that) do not typically refer to a specific referent”. But Cheyney and Seyfarth have shown referential alarm calls in vervet monkeys, and songbirds have aerial and terrestrial alarm calls that differ in frequency characteristics. There are many (review) papers on referential alarm calling. Can you clarify what you mean, as you elicited alarm calling with a model snake, falcon, and currawong? Perhaps the model birds were on the ground, which is known to elicit terrestrial rather than aerial alarm calls.

Abstract

I found the abstract rather choppy. In line 3 you refer to “similar vocalisations”, but one only understands the meaning after reading the manuscript. In line 5 you write “this vocalisation globally”, which is again confusing. I think you mean over large geographical areas. In line 6 you write “this vocalisation is functionally referential”, but you mean “this vocalisation type” (referring to the whining call type). Your repeated reference to “this vocalisation” makes it seem as if there is one vocalisation, but you mean the whining call type with similar frequency characteristics across geographical areas. I know space is limited -- but the abstract was hard work. Finally, in line 14, you write “is learned for its production”, and again, if possible, clarify “sound production”. The last sentence would benefit from adding “of sound production” at the end.

Line 52: the term innate vocal reactions appears for the first time, and then throughout the manuscript. I think I understand what you mean, but what is a little difficult to quickly grasp is the fact that the “reactions” you refer to are, as I understand it, reactions to internal states or external triggers that result in innate vocal production. So innate vocal reaction is equivalent to innate vocal production (as the consequence of a reaction to some cue or state). Is there any sense to make this point here?

Line 74: usually we think of selection on honest signalling, not on honest responding. Maybe delete “honest” before “responding”? The predictability is important for the statistical probability that may underpin forms of associative learning.

Line 87: might be worth noting somewhere that you are referring to vocal production learning systems. The complication with the terminology is that the so-called “vocal non-production learning systems” can have associative sound learning.

Line 92: you write “but their response to it is innate”. Is the vocal production also innate?

Line 96: replace word “using” with “producing” a similar sound solution

Lines 104-106: This issue of alarm calls is the biggest weakness of the manuscript. Clearly, birds and mammals have referential alarm calls. In songbirds, aerial alarm calls tend to be whistles, and terrestrial alarm calls tend to be broad frequency bandwidth calls, where urgency is also communicated by call rate.

Line 116: which predator types and hence which predator-associated alarm call types?

Line 126 and 133: which predator-associated alarm call types?

Line 138: innate danger? Perhaps rephrase as innate vocal response to a threat.

Figure 1: there is no ‘general alarm call’ that I am aware of. There are terrestrial and aerial alarm calls, for example.

Line 141: insert word 'host' before defenses

Line 161: why global? I think you mean "with the geographical distance of species"

Line 170: replace "globally" with "at larger geographical scales", or something like that.

Lines 188 and 194 etc – the general alarm call issue.

Line 242: insert "to whining calls" after "other species responded"

Line 321: when you refer to the median number of vocalizations per species – do you mean host species or predator species?

Lawson et al 2023 in Behavioural Ecology and Sociobiology showed that Galapagos Yellow Warblers do not produce the anti-brood parasitic seet call, nor do they respond to an experimental broadcast of the North American Yellow Warbler's seet call. The authors conclude that this supports the view that the association between the seet call and a brood parasite is learned, as you write. How does this island-endemic response pattern fit with your findings? This might be interesting to consider, as the Galapagos Yellow Warblers are thought to have diverged recently, within the last 300,000 years. Or perhaps the whining call is a special class of calls?

Can you provide spectrograms of the various call types across geographical areas, including examples of the alarm calls for the experiments?

Reviewer #5 (Remarks to the Author):

This paper examines the occurrence of a brood parasite specific alarm call given by hosts of brood parasites (the whine call) and frames the study as important in terms of the evolution of human language. The authors first point out the features of Darwin's idea about language evolution and then point out the possible connections with the alarm call of cuckoo hosts. They then explore several aspects of the whining call. First, they show that the whine calls are taxonomically widespread, that these calls given by various hosts of brood parasites are all similar in structure to each other but differ from alarm calls given to predators. Second, they do a comparative (macroevolutionary) study that shows that the whining call is predicted not by cooperative breeding (and hence communication within groups) but by the richness and complexity of the brood parasitic community a give host lives in (which they speculate indicates that calls function in heterospecific communication among hosts). Third, playback experiments combined with model presentations provided evidence that a second species (scrubwren) uses whining calls in the context of brood parasitism (previously shown for fairy-wren). Finally, playbacks showed that species respond appropriately to whine calls even from taxa they are allopatric with

In terms of positives, this paper explores a very interesting phenomenon, that hosts of brood parasites appear to use an alarm call that is specific to brood parasites and that this call may have evolved early in the songbirds or independently in various taxa. There is some evidence for this claim, but also some clarification or additional information is needed for this to be fully convincing. I found the paper a bit of struggle to read and to see how the parts fit together. Some of this could be because the initial framing of the study adds an unnecessary complication in terms of having the authors try to tick the boxes for that framing rather than focusing entirely on making a convincing case for the use of specific alarm calls aimed at cuckoos. To be fair to the authors and make sure that I was not being overly crabby about the paper I shared it with a colleague (name provided to the editor) who is an expert in alarm calls and mobbing behavior in birds. That colleague had a very similar response, both overall and to specific points I outline below. I do not believe the paper is publishable in its current form and it needs to be considerably reframed, and certain aspects need to be made clearer with additional details.

1. In my opinion, the framing of the paper as presenting an example that supports Darwin's musings about language evolution does not work. This framing made the study harder to follow because the reader is focused on ensuring that the assumptions that would support Darwin are met, rather than focusing on the topic of a specialized brood parasite alarm call itself, which is interesting in its own right. This framing makes the paper a bit hard to follow in places and gets in the way of clarity. In addition, the alarm call seems like a bad fit with the human language framing because human words are learned, whereas the alarm calls are innate. The fact that these calls 'sit between the extremes of innate and learned' (line 90) is irrelevant. The fact that the context is learned, but not the sounds themselves, means the connection to human language does not work.

2. A concern for me was that I was not fully able to consider the results of this study in the context of the literature on mobbing vocalizations and to be convinced that what is shown here is clearly distinct. There are a few concerns. First, the way the hosts respond to cuckoos seem like a mobbing reaction and the vocalizations are somewhat similar in structure to mobbing calls. Despite this similarity, the term mobbing is never used in the paper and possible connections to predator mobbing are not discussed. Second, it is well known that many songbirds (at least temperate zone birds) have two distinct types of alarm calls, a pure tone hard to locate seet call that signifies an immediate threat like a flying raptor, and a more broadband with harmonic structure call that is used in mobbing, is easily to locate for recruiting other mobbers. The authors lump all predator alarm calls together and it is hard to know if how much of the difference between whines and alarm calls is

because they have used mainly seet type calls. For example at line 104 "To test this, we collated and analyzed the acoustic structure of vocalizations produced by hosts towards brood parasites and predators, as predators also present a threat but are typically met with alarm vocalizations that do not typically refer to a specific referent (hereafter 'alarm' vocalizations)". I think using mobbing type calls and not seet type calls would be more appropriate for this comparison. Note that if seet calls were compared to mobbing calls in many songbirds they would show up as strikingly different, more so that the differences shown in the current paper where it was the third level of the PC (PC3) where separation was shown. Perhaps the distinction I am making here (mobbing vs seet calls) does not apply to some of the taxa investigated here but these details should at least be considered and discussed.

2. Following on the point above about mobbing calls, given that the context to produce the whining calls is learned, I was surprised not to see any reference to Curio's famous work on this topic for alarm calls generally. In an elegant series of lab experiments Curio showed convincingly that some young songbirds learn to associate alarm calls with a particular stimulus they are seeing when they hear other birds alarming. Songbirds were taught that honeyeaters are dangerous, and that dish detergent bottles are also dangerous, and if I recall correctly these specific patterns of enemy recognition were transmitted across future generations (young birds imprinting on harmless items then passed this on to their own offspring). This work is relevant to the subject of this study and may also be relevant to the issue of how unique the brood parasitic alarm calls are (or whether they are just a slightly modified version of more general mobbing calls. In some cases, might they be more generally mobbing calls that become associated with parasitism through learning? In this context is also interesting that some cuckoos mimic birds of prey (Welbergen and Davies) and I expect are mobbed as such.

Minor points by line numbers

43. Is this really analogous to language? If so, are all alarm calls and mobbing calls similar?

53. Jargon. What does innate vocal reaction mean? What are some examples.

56. If it is arbitrary then why would we see convergence. I do not think this makes sense.

I do not know about linguistics and ideas about human language to be able to make sense of these claims.

69. There is a mix of things here that is then summarized as a suggestion that does not make sense to me. Some aspects of words are shared (r in round) in human speech AND human screams and distress vocalizations elicit an innate recognition. Putting these two completely different things together " suggest that some aspects of behavioral responses in an otherwise learned system harness phylogenetically generalizable innate associations between sounds and meanings, perhaps due to them delivering an adaptive benefit". I do not understand how these things combine to lead to the suggestion.

101. (already summarizes above but included here for clarity) I am very surprised that nothing is mentioned about mobbing because in some ways these calls seem structurally similar to mobbing calls. It is also interesting that some cuckoos mimic raptors and are likely to mimic a mobbing response. If these whining calls are in some ways related to mobbing calls or derived from them, then the story is somewhat different than presented.

If whining calls are basically mobbing calls, or are even modified mobbing calls, then the notion that whining calls are "functionally referential and harness an innate bias between sound and meaning" does not seem totally correct to me. That statement implies that all of these patterns and effects are exclusive to the specific context of brood parasitism. The effects of enemy recognition more broadly (predators) which is widespread among birds, needs to be parsed out.

135. "...contain conserved acoustic features that are known to elicit an innate danger response in other species". I am not sure what is being said here.

156. Were the calls that were given to cuckoos versus predators analyzed as was done earlier in the paper to confirm that these birds are producing calls distinct for brood parasitism and that these differ from mobbing calls given to predators?

A couple of the sounds from Xeno Canto seem potentially problematic. For *Prinia rufescens* there were several recordings of birds giving alarm calls and many sounded very different from each other. How did the authors decide to choose the example they chose. And for the *Philemon corniculatus* example, the bird seems to be giving a regular call (and there is no mention of alarm call in the recording description).

Version 1:

Decision Letter:

24th June 2025

*Please ensure you delete the link to your author homepage in this e-mail if you wish to forward it to your co-authors.

Dear Will

Your manuscript entitled "Learned use of an innate sound-meaning association in birds" has been seen again by the reviewers, whose comments are attached. It's clear that the latest version is considerably improved, and two reviewers now

sign off. Reviewers 2 and 5 have some further questions and comments that will need to be addressed, and we have also had to add another reviewer, reviewer 6, to cover the expertise of reviewer 1 who was unable to re-review. Reviewer 1 was invited partially due to their phylogenetics expertise, though they didn't make any comments in that regard in the first round--I had intended to follow up with them whether they had any specific concerns about those aspects of the manuscript this round, so when they couldn't re-review we felt it best to get another reviewer with that expertise whose comments are now attached. I realise that in some ways it will be frustrating to get extensive additional comments at this stage in the process, but the most important thing is that all the technical details are ironed out especially in a paper of this kind which has so many different components. With that in mind, please could you pay diligent attention to the reports from reviewers 2,5, and 6, in your next revision.

We therefore invite you to revise your manuscript taking into account all reviewer and editor comments. Please highlight all changes in the manuscript text file [OPTIONAL: in Microsoft Word format].

We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact us if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

When revising your manuscript:

- * Include a "Response to reviewers" document detailing, point-by-point, how you addressed each reviewer comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the reviewers along with the revised manuscript.

- * If you have not done so already please begin to revise your manuscript so that it conforms to our Article format instructions at <http://www.nature.com/natecolevol/info/final-submission>. Refer also to any guidelines provided in this letter.

- * Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

- * Extended Data Figures - please ensure that any supplementary figures and tables that are crucial to the manuscript's conclusions are converted into Extended Data figures and tables to increase visibility of these data. Extended Data figures and tables are online-only (present in the online PDF and full-text HTML versions of the paper), peer-reviewed display items that provide essential background to the article but are not included in the main article due to space constraints. A maximum of ten Extended Data display items (figures and tables) is permitted.

Please use the link below to submit your revised manuscript and related files:

Link Redacted

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

We hope to receive your revised manuscript within four to eight weeks. If you cannot send it within this time, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Ecology & Evolution or published elsewhere.

Nature Ecology & Evolution is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit <http://www.springernature.com/orcid>.

Please do not hesitate to contact me if you have any questions or would like to discuss these revisions further.

We look forward to seeing the revised manuscript and thank you for the opportunity to review your work.

Yours sincerely,

[redacted]

Reviewer expertise:

as before, reviewer 6 has expertise in avian macroevolution and phylogenetics

Reviewers' comments:

Reviewer #2 (Remarks to the Author):

I think that the additional detail provided by the authors strengthens the argument that whining calls are functionally referential to brood parasites and learned, which was my primary concern in my initial review. I do still have some minor concerns listed below, but if the authors can address these then I recommend that the paper be published. I think it is a fascinating and important study and would like to see it get the attention it deserves in a high-impact journal like *Nature Ecology and Evolution*.

1. I appreciate the addition of an analysis to assess if birds exhibit a qualitatively different response to whining calls vs. terrestrial alarm calls. However, the response variable chosen (number of individuals approaching 0-2 meters vs. 2-5 meters) seems a little contrived to me, as if you were trying to artificially convert a continuous variable into a categorical response. For many of the host species you tested, I understand that there may not have been any other clear response to measure. However, in the case of yellow warblers, why not record whether or not the warblers flew back to their nest in response to the playbacks? That seems like the most obvious response variable to measure, since it apparently represents a clear qualitative difference in how yellow warblers respond to brood parasites vs. predators (line 154). Was there a reason you did not include this response variable?

2. The description of the results for the playback experiment with yellow warblers is a little unclear. The wording of lines 542-544 makes it sound like the lack of a differential response to the treatments in yellow warblers contradicts your hypothesis that whining calls elicit a universal innate response in passerines. But if I understand correctly, the fact that yellow warblers responded equally to fairy wren whining calls and to yellow warbler seet calls actually supports the hypothesis that responses to whining calls are innate and universal. Can you please clarify in the text whether the responses of yellow warblers to the playbacks supports or contradicts your hypothesis?

3. In Figure S2 it looks like yellow warblers responded equally strongly to whining, chip, and seet playbacks, but much less strongly to control playbacks, yet line 543 of the main manuscript seems to indicate that there were no differences in response to any of these treatments. Can you clarify if indeed there was no statistically significant difference in response between the control playbacks and the other playbacks?

4. I think you should provide a little more clarification about what you mean when you say that the response to whining calls is innate. Yes, you found that whining calls elicit approach from passerines with no prior exposure to whining calls. But this doesn't necessarily mean that these species understand whining calls as specific references to brood parasites. Rather, whining calls may simply elicit a heightened but non-specific response in naïve species by taking advantage of a pre-existing auditory bias in passerines. To determine if naïve species understand whining calls as referential labels for brood parasites, you would need to do some additional experiments. For example, if yellow warblers fly back to their nest in response to both seet calls and whining calls but not in response to chip calls, then this would imply that they innately treat whining calls as referential signals for brood parasites (which would be truly astounding!). Similarly, if naïve species respond to whining call playbacks in the breeding season but not outside the breeding season, that might imply that they innately understand the calls as references to a threat that is only relevant during the breeding season (although you would need a control to ensure that the birds are not simply less responsive to all warning signals outside of the breeding season). I don't expect you to do these additional experiments for the current study, but I think you should provide some clarification in the manuscript about exactly what you mean when you say that the calls elicit an innate response. If naïve species are innately sensitive to the acoustic properties of whining calls but don't understand them as referential signals, would these calls still represent a bridge between innate-reactionary and learned-symbolic signals?

Reviewer #2 (Remarks on code availability):

Unfortunately, the link to the Dryad data/code still does not work for me. When I clicked on the link I got "Error: DOI not found".

Reviewer #3 (Remarks to the Author):

The authors properly addressed all my comments and the MS is now much improved both in clarity and in statistical robustness. I have no further comments.

Reviewer #4 (Remarks to the Author):

I thank the authors for this substantial and thoughtful revision. They have fully addressed my previous comments, and their responses to the other reviewers' concerns are likewise satisfactory.

The study provides good evidence for a previously undocumented link between innate and learned vocalizations, representing a kind of missing link in the evolutionary history of language. The tantalizing implications for the evolution of communication systems, including the foundations of language, are further strengthened by the cross-continental scope of the experimental work, which significantly enhances the generality and impact of the findings.

I believe Darwin would have been most pleased with this elegant demonstration of how innate and learned elements may

co-vary in communication systems.

Sonia Kleindorfer

Reviewer #5 (Remarks to the Author):

The authors have done a very thorough job of responding to all of the reviewers' comments and it is clear that they have an extremely interesting study with a wealth of experimental and observational lines of evidence to support their claims. I also like the very clear organization, with subsequent sections each clearly addressing a key question or component.

That said, I am still not sure that the authors have completely addressed a key concern I raised, a concern also raised by other reviewers. The issue is whether this whining call is distinct from the general mobbing calls that many species use to recruit other birds to mob a predator. At least in northern hemisphere songbirds, there are two classic "alarm" calls—a high pitched pure tone "seet" call that is used in response to a flying raptor. This results in birds freezing or diving to cover and is not associated with mobbing. These calls are also hard to locate. All of this makes sense because the predator poses an immediate threat to the lives of the birds giving the alarm call. In contrast, classic mobbing calls are broad frequency, often harmonic calls that are used to recruit other species to mob a predator. The structure of these calls makes them easy to locate—this helps in other mobbers join in the mobbing of the predator. These calls are used when there is a perched predators (or snake on the ground), which are much less threatening to the mobbers' survival than is flying predator. In some cases birds approach the model but are not aggressive towards it; in other cases, mobbers will actually attack the predator (hit the raptor on the head from behind). This is relevant the authors' point that hosts will not physically attack some birds (currawongs) but will attack others (cuckoos). This distinction (attack versus just come close) may not be useful to distinguish among types of alarm calls.

With this in mind, I am still a bit confused about the terminology the authors use and I am not convinced they have fully dealt with this issue in their response to reviewers. They refer to "terrestrial alarms". Are these used to recruit mobbers or do they function like the high frequency seet calls I mentioned above: the birds freeze or hide? Put another way, for the authors' claims to be correct, the birds with whining calls would also need to have a second call that also recruits mobbers, but for something other than a brood parasite. Alternatively, if the second alarm call functions like the seet alarm call then the claims made by the authors are overstated. In that case, the whine call of the hosts could simply be a classic mobbing call that is learned to be used towards cuckoo, and their only real mobbing call. (As Curio's experiments showed, birds could learn that this mobbing call is for cuckoos and not other birds like owls or sitting raptors). In this case, we would not be talking about the evolution of a new alarm call type (whining) but simply using the mobbing type calls that many songbirds have for a specific enemy. And if these are just typical mobbing calls, and if these hosts do not have other mobbing calls but more high threat alarm calls, then it would not be remarkable that they differ in structure from the other alarm calls. An analysis of mobbing versus seet calls would show exactly the same thing. I raised this point in my initial review but despite the voluminous response from the authors to me and the other reviewers who raised similar issues, I am still not sure that I have a clear understanding of what the terms alarm calls and terrestrial alarm calls. Perhaps I am missing something. If the authors have clearly identified a call (whine) that is different and distinct from the more classic mobbing call (i.e. call used to recruit other birds to mob a sitting predator of some sort (typically bird but also snakes in some birds) while at the same time these species have a second different mobbing call for sitting predators that is used to recruit mobbers (i.e. not a call to indicate an immediate threat that is not used to recruit mobbers, then I think the paper is very novel and should be published. However, if these species have only one call that recruits mobbers, and this is used for cuckoos, and they lack a second different call to recruit mobbers for other threats, then the paper is not exactly what it claims to be. I ask that the authors provide a brief clear response to this question so that I can fully understand what they are showing here and what is unique and novel.

One thing that may be confusing me, and this study in general, is that northern and southern hemisphere birds may have different sound structures and uses of alarm calls. If this were the case, using the same labels for all of these species. Note that the prinia and the Old World warbler both have what looks like a pure tone "alarm" vocalization that may function as the standard "flying predator warning" rather than a recruitment for mobbers.

Reviewer #6 (Remarks to the Author):

The authors use a combination of experimental and modelling approaches to demonstrate how learned vocalisations may derive from innate responses with shared evolutionary origin. The overall idea is fascinating and may be of interest to evolutionary biologists and linguistics. However, the fairly small sample sizes and inaccurate descriptions of statistical methods and results throughout the manuscript (see comments below) raise concerns about the scientific rigour of the study. The authors will need to improve their analyses and be much more transparent with their methods and results before the manuscript could be considered for publication.

Major concern:

One main conclusion of this ms is that whether a species produces whining calls (presence or absence) is associated with opportunities of inter-species cooperation (measured as linkage density). The authors have tried to argue that the two measures of cause and consequence are independent since data were collected from different sources (call type mainly from recordings and literature, and linkage density solely from literature), but the positive association may also be caused by the biases in previous knowledge. Common/better-studied species tend to have more complete information on parasite-host interactions and thus higher linkage density, have more xeno-canto recordings which make whining calls more detectable,

and also are more likely to be documented as producing whining calls in literature – the last can be very problematic since nearly half of the species were assigned their whining status (yes or no) based on literature alone. If the authors can repeat their models by limiting the samples to well-known species in well-studied areas (e.g. in North America/ Europe/ Australia) to reduce sampling bias and find similar results, the conclusion may be more convincing.

The following points are sort of minor but show a lack of understanding of statistical methods and/or sloppy writing.

Line 184-185: '> 5 times more slowly in whining than in alarm vocalizations'? Is it alarm vocalisations or all vocalisation types put together? The LRT result suggested the latter. And what is 'LRT comparing models with and without vocalization-specific evolutionary rates' anyway? There is no text about LRT models in methods so it's hard to tell. Please add relevant model details. And what does PC3 mean in line 185 of the LRT result?

Line 427-428: 'phylogeny explained much of the variation in species that produce whining vocalizations' – what is 'variation in species'? If you meant whether or not (a binary state) a species produces whining calls then please write it that way. Also, Pagle's lambda is not equal to how much variation explained by phylogeny (although this is common misconception), but only to what extent related species share similar traits to each other. So it's better to say 'producing whining vocalisations is phylogenetically conserved' or something similar.

Line 432: I don't think Fig. 1C shows anything about within-species cooperation.

Line 434-437: chaotic presentation of results. What do Bayesian model results say about the geographical distribution of whining species? I think the models only test whether or not a species produces whining calls. In line 437, how could Fig. 1C support model results of how many trees were supported? I understand the figure may corroborate the results in a way but citing a figure directly after phylogenetic model results suggests they convey exactly the same information.

Line 438-442: In my view Fig. 1c is probably more suitable for supporting these results, only.

In methods, when you refer to R functions please use quotation marks around the function names (or italicise) and give citation or a link to every package. For example, line 655 'we conducted a principal components analysis using princomp on all species and vocalizations' is poorly written. I know that princomp is a basic R function for PCA but writing it as if it was a regular word is inappropriate. Please change all relevant cases throughout the ms.

Line 683: 'observed and expect counts' of what vocalisation type, whining or alarm? I assumed the former given the context but then it contradicts 'greater than 2 was considered... less often than expected as whining vocalisations'. Is it actually less often or more often?

Line 689: linear mixed effects models, not 'linear effects models'.

Line 694: vague method, unclear how the whining status was defined using emmeans. E.g., 'differed in post-hoc tests', differed in terms of different vocalization contexts (brood parasite-associated vs alarm)? I couldn't tell. If the linear discriminant scores 'capture the acoustic features of vocalizations along the axis separating parasite-associated and alarm vocalizations', the scores must be higher or lower (i.e., uni-directional) for a species to be considered whining, not just being different. From your responses to other reviewers, it sounded like you also included an interaction term between species and calling context, this is crucial for understanding the method but it is not mentioned anywhere in methods. Please specify the methods.

Line 768: What was the prior used? What was the thinning rate and the final sample size of models? These are important for Bayesian models and I couldn't find any information. The model results are also presented very vaguely. They are only mentioned once in-line without much detail. A table detailing the sample size, effect size and CI of predictors, etc. is needed.

Line 753: 'As host status for a given species may change through time, we only use documented host species in this analysis' – I don't understand what it means.

Other comments:

Abstract: '..., indicating that convergent use facilitates cooperative defenses'. Adding 'across species' at the end may improve clarity given it is said just a few lines above 'species which produce this vocalization are rarely socially cooperative'.

Figure 1A. The caption only mentioned species with and without whining calls identified (n = ~50), which are far less than the number of tips and families shown in the phylogenetic tree. I assume the tree shows all passerine families and the tips are all passerine species (n = ~5k)? Please explain what the tree consists of, and add sample sizes for how many species are identified to produce whining/seet calls, how many are not. There also seems to be a purple/grey bar above the warbler which is different from whining/seet/no unique vocalisation, what call type is it? The caption of fig. 1C is also very uninformative about how the linkage density was calculated. Just citing one reference is not enough, please add details. Figure 2. what species are involved in the panel K? Not explained in the main text or caption. Perhaps it's not key but without such information it's impossible to read a phylogram. A version with all species names added needs to appear in supplementary information at least.

Line 690: adding sample size for glmmBMT models will improve transparency and credibility of results (e.g., what was the total number of vocalisations/scores? Should be 25 species × ~40 per species?)

Line 756-757: Kennerley et al.25 should be Kennerley et al.22? Check references.

Reviewer #6 (Remarks on code availability):

The link to data and code is still not working. DOI not found.

*****END*****

Version 2:

Decision Letter:

31st July 2025

Dear Dr. Feeney,

I am writing in the temporary absence of my colleague, Dr. Luíseach Nic Eoin.

Thank you for submitting your revised manuscript "Learned use of an innate sound-meaning association in birds" (NATECOLEVOL-24102745B). It has now been seen again by the original reviewers and their comments are below. The reviewers find that the paper has improved in revision, and therefore we'll be happy in principle to publish it in Nature Ecology & Evolution, pending minor revisions to satisfy the reviewers' final requests and to comply with our editorial and formatting guidelines.

Please email us a copy of the file in an editable format (Microsoft Word or LaTeX)-- we can not proceed with PDFs at this stage.

Please also email us a completed copy of the Reporting summary :

Reporting summary: https://www.nature.com/documents/nr-reporting-summary.pdf

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements in about a week. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Thank you again for your interest in Nature Ecology & Evolution. Please do not hesitate to contact me if you have any questions.

Sincerely,

[redacted]

Reviewer #2 (Remarks to the Author):

The authors have addressed all my concerns, and I now recommend that this article be published. I only have one very minor proof-reading comment:

In line 770, when you say "we THEN produced a hexgrid", it makes it sound like producing the hexgrid is a separate step from the previous sentence that talked about calculating the opportunity for between-species competition. But in fact, as far as I can tell, producing the hexgrid was part of the process of calculating the opportunity for between-species competition. So I suggest removing the word "then" in line 770.

Reviewer #2 (Remarks on code availability):

The link to the code on Dryad still does not work. I have never been able to access the code throughout this review process.

Reviewer #5 (Remarks to the Author):

I thank the authors for carefully addressing my remaining questions. I still have slight lingering doubts about whether the whining call is unique as opposed to being a more general mobbing recruitment call that has been learned to apply specifically to cuckoos. Perhaps that is a non-issue or is consistent with what the authors are claiming. Given the number of reviewers on this paper, if this were a big concern others would have raised it as well. Given this, I am happy to sign off on this for publication. It is a really interesting paper and it will stimulate further work on the topic.

Reviewer #6 (Remarks to the Author):

The authors have addressed all of my comments and have much improved the clarity of their result presentation, methods, and manuscript text. I appreciate their effort of revising the manuscript and recommend its publication.

Minor edit: in response to my previous comments, lines 445-447 were added to justify why knowledge bias may not affect result robustness. The author provided good logical argument in their response but the rationale was not fully explained in the ms, so it may be clearer to add some points to show readers the contrast between research effort and documented linkage density.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Reviewer #1 (Remarks to the Author):

My main concern in terms of the quality of evidence supporting the main claims was as follows. The whining call, in my experience, is not limited to the context of brood parasites. It seems to be used more generally as a warning signal in the context of risks to the nest contents (rather than to adults). It may be elicited by predators such as snakes or rats that pose minimal risk to the parent birds, or, in some bird species, even by people closely approaching nest sites. I would question whether these vocalisations are referential to brood parasites alone. The manuscript does attempt to fend off this type of criticism with an experimental test involving presentation of a supposed nest-predator (snake) near active nests. While this treatment did not seem to elicit whining notes, the sample sizes are not great and I am not sure that an immobile snake model elicits the same response as the more slithering variety.

RESPONSE: We hope that our response reassures the reviewer that all aspects of our study design and approach are consistent with published literature, and that our analyses are based on assumptions that have been checked and met. As this comment includes multiple parts, we have provided separate responses to each.

Regarding the targeted use of the whining vocalizations in a specific context: Our understanding is that whining vocalizations were first described in the splendid fairy-wren (*Malurus splendens*) in response to a model cuckoo (Payne et al. 1985 Behav. 94: 108-126), which included an examination of whether it was made towards cuckoos when they did and did not present a threat (e.g. eggs or chicks in the nest vs after the chicks had fledged). Those authors found that whining vocalizations were exclusively made towards cuckoos when they did present a threat (i.e. eggs or chicks in the nest) but not made when cuckoos did not present a threat. Since then, these vocalizations have been described in the superb fairy-wren (*Malurus cyaneus*), with experimental studies finding that its production in the correct context (towards cuckoos) is socially learned despite hearing it eliciting an innate rapid recruiting response (Langmore et al. 2012 Behav. Ecol. 23: 798-805; Feeney & Langmore 2013 Biol. Lett. 9: 20130443), and that its production towards cuckoos is functionally referential, i.e. only being made towards cuckoos and hearing it eliciting a distinct behavioural response compared to responses towards other vocalizations in their repertoire (Feeney et al. 2013 Science 342: 1506-1508). Moreover, despite fairy-wrens being the focus of numerous long-term field-based studies, including many studies examining behavioural responses towards various nest threats (e.g. Magrath et al. 2007 Behav. Ecol. 18: 944-951; Grieg & Pruett-Jones 2009 Anim. Behav. 78: 45-52; Colombelli-Négrel et al. 2010 Behav. 87: 853-870; Teunissen et al. 2021 Curr. Biol. 31: 4120-4126), we know of no other studies that have examined or described this vocalization,

nor any that have suggested that whining vocalizations are made in response to anything other than cuckoos, or towards cuckoos when they do not present a threat to a host nest. Apart from fairy-wrens, strikingly similar vocalizations have been reported in the large-billed gerygone (*Gerygone magnirostris*) towards their cuckoo in north-eastern Australia (Noh et al. 2021 Anim. Behav. 174: 51-61), as well as in numerous old world leaf warblers (Phylloscopidae) towards their cuckoo throughout Eurasia (Wheatcroft and Price 2015 Behav. Ecol. 26: 83-90). The data presented in this present manuscript build on these past studies by providing acoustic analyses indicating that 21 (of 26) host species produce whining vocalizations towards brood parasites, a literature analysis that found no evidence that these vocalizations are used towards anything other than brood parasites, and experiments on multiple species in Australia and China, which found that whining vocalizations are used towards brood parasites (including that it is functionally referential in an additional species: the white-browed scrubwren, *Sericornis frontalis*).

We appreciate this comment as it highlights that we could have been clearer regarding the documentation of whining vocalizations in the literature, and that they have only been reported towards brood parasites and only when brood parasites present a threat. To address this, we have now added a more detailed discussion of past literature in the introduction, and also edited the discussion to highlight that whining and distress vocalizations have previously been examined and shown to be structurally distinct.

Regarding the ecological relevance of the snake treatment, whether birds respond appropriately towards it, and our statistical approach: Eastern brown snakes are known predators of superb fairy-wren nests (Yasukawa & Cockburn 2009 The Auk 126: 147-154) and using mounts of them to elicit a response comprises an established experimental method (e.g. Feeney et al. 2013 Science 342: 1506-1508). Regarding the hosts' response we recorded to them, our supplementary data shows that the snake treatment did elicit an ecologically appropriate response, including the production of (terrestrial) alarm vocalizations and a physical response that resembled responses towards the other predator treatments in both superb fairy-wrens and white-browed scrubwrens (Table 1). We have not made any changes to the text, because the response to the snake, while interesting, is only important in this study as a comparison with the cuckoo treatment. Regarding sample sizes, ours are comparable to (or larger than) those in other brood parasitism studies that use model presentation experiments (e.g. Davies & Welbergen 2009 Science 324: 1318-1320; Feeney et al. 2013 Science 342: 1506-1508). Additionally, we note that we have provided an R Script that includes assumption tests for all statistical analyses that we conducted (see Supplementary R Code) and that assumptions were met for all analyses presented in this manuscript.

	Superb fairy-wren	White-browed scrubwren
Treatment	Number of terrestrial alarm vocalizations	
Snake	86.53 $\pm$ 27.98	113.18 $\pm$ 47.83
Raptor	169.47 $\pm$ 34.58	126.76 $\pm$ 44.16
Currawong	68.53 $\pm$ 32.46	173.12 $\pm$ 65.32
	Time spent by at least one bird between 0.5-5m of model	
Snake	196.06 $\pm$ 23.95	207.88 $\pm$ 24.90
Raptor	189.94 $\pm$ 26.66	205.65 $\pm$ 23.03
Currawong	193.59 $\pm$ 27.05	186 $\pm$ 26.86

Table 1. Number of alarm vocalizations and time spent by at least one host bird between 0.5m - 5m of the model during the snake, raptor and currawong (predator) treatments of our model presentation experiments (n = 17 for all treatments) (mean $\pm$ standard error). Alarm vocalizations are presented as they are the primary vocal response towards predators, and 0.5-5m is presented as it is where the birds from both species spent most of their time during the predator treatments (see Supplementary Data for full dataset).

If the whining call type is a more general signal of risk to nests or nest-contents then it seems a bit harder to argue for direct symbolic association between the vocal signal and the context of brood parasitism. Both nest-predators and brood parasites can be mobbed by conspecifics (and heterospecifics), so a more general function of the whining call to communicate risks is possible. It is necessary to consider whether we, as humans, may be superimposing the meaning of ‘brood-parasite’ onto the sound, whereas the organism is using the call as a catch-all signal to convey a variety of risks to nest contents. Brood parasites are then potentially batched together with nest predators in the sound symbolising “risk” and species are able to learn new risks, be they from parasites or predators. If the innate response to nest-risk signals is universal, then the process of learning which species pose risks – whether through predation or parasitism –can also be viewed as equally universal, rather than convergent as claimed in the paper. I think the text needs some editing to add more caution to the conclusions, and to acknowledge that the function of the whining signal requires further testing in a broader array of host species, and with more a convincing experimental set-up, including treatment with a larger sample of possible nest-predators.

RESPONSE: Following our response to the previous comment, we know of no published literature that suggests the whining vocalization is a general signal of risk. All published literature that we know of, as well as the data presented in this manuscript, indicates that the whining vocalization is a functionally referential signal for a brood parasitic cuckoo. Regarding our experimental approach, the spectrum of threats we used include a major predator from each threat category that these birds face (i.e. predator that primarily targets adult birds, predator that primarily targets nest contents, generalist predator, brood parasite and a control), which follows an established approach (e.g. Feeney et al. 2013 *Science* 342: 1506-1508). We also note that despite numerous other studies presenting an array of fairy-wren species with various other threats (e.g. Magrath et al. 2007 *Behav. Ecol.* 18: 944-951; Grieg & Pruett-Jones 2009 *Anim. Behav.* 78: 45-52; Colombelli-Négrel et al. 2010 *Behav.* 87: 853-870; Teunissen et al. 2021 *Curr. Biol.* 31: 4120-4126), “whining” vocalizations have only been reported towards cuckoos (and only when cuckoos present a threat to the host nest).

Another concern is the assumptions around the category of non-hosts. Interpretation of statistical differences between current hosts and non-hosts is potentially problematic because the identity of host species undergoes turnover through time. Brood parasites move from host to host depending on arms-race dynamics. That is, when one host develops strong anti-parasite defences the brood-parasite benefits from switching to a naïve host. One upshot of this is that what may seem like a non-host may well have been a host in its recent evolutionary history, therefore retaining innate or learned responses to brood parasites and warning signals from conspecifics or heterospecifics (the ghost of parasitism past). In addition, the distinction between hosts and non-hosts may be blurred by biases in knowledge, with some hosts overlooked in data-poor cases. Some discussion of the potential pitfalls of inferring or assuming non-host status on the basis of current information is warranted.

RESPONSE: In this study, we use documented instances of host use by a brood parasite to inform our classification of host status, which is the same as and follows the approach used by other studies on the topic (e.g. Feeney et al. 2013 *Science* 342: 1506-1508; Medina et al. 2016 *Ecology Letters* 19: 1110-1118; Antonson et al. 2020 *Nature Communications* 11: 4185; Kennerley et al. 2022 *Ecology Letters* 25: 1889-1904). Importantly, this approach is highly conservative (i.e. any undocumented hosts would drag our results towards being non-significant), and given this, that we found significant results highlights the strength of the signal in the data. Nonetheless, we agree that more information would be helpful here, and we have expanded the methods text accordingly.

For the reasons above, I think the authors need to dial down the level of certainty they ascribe to the findings and conclusions. Despite these shortcomings, the overall message of this manuscript remains potentially interesting. The analyses provide incomplete but suggestive evidence that

signals eliciting innate responses can be used as referential symbols, and applied to learned objects. The suggestion that this type of scenario could provide early building blocks for the development of more complex language is intriguing, but highly speculative. Whether the finding is as novel as the authors state needs to be checked by experts in related fields, including linguistics.

RESPONSE: Having now modified the text following the suggestions of all reviewers, we are confident that our manuscript has been greatly improved. We appreciate this reviewer's positive appraisal of how our findings fit in the context of the broader literature, and note that the senior author on our paper is a linguist whose previous research includes the seminal finding that sound-meaning associations are a generalizable feature across human languages (Blasi et al. 2016 PNAS 113: 10818-10823).

Reviewer #2 (Remarks to the Author):

I commend the authors on an incredibly comprehensive study in which field experiments conducted at multiple sites around the world were combined with acoustic and phylogenetic analyses to address an important evolutionary question. The question of how sound-meaning associations is likely to be of great interest to scientists from a variety of disciplines, including behavioral ecologists, evolutionary biologists, and linguists. This study attempts to address this question in a way that few others have by quantitatively investigating the evolution of referential calls. The manuscript is very well-written and lays out a compelling rationale for the study as well as clearly explaining the experiments and analyses that were conducted. The experiments were well designed and included proper controls (although I am not sure if the response measured was the most appropriate for the research question- see below). The figures are extremely well done and informative. I especially like Figure 1. I cannot comment on the methodology for the phylogenetic and linkage density analyses as this is outside my area of expertise.

However, despite my overall positive opinion of the paper I do have a concern about the central claim that whining calls are functionally referential and learned. The authors argue that their findings are especially novel and important because there are few if any previous examples of functionally referential calls with learned production and an innate response, and finding an example of such calls would support the hypothesis that symbolic meaning arises from pre-existing innate associations. But based on the evidence provided in this paper, I am not completely convinced that whining calls are functionally referential or learned at all.

To be considered functionally referential, a call must refer to a specific external stimulus, rather than simply being a vocal manifestation of the caller's internal emotional state. To demonstrate that a call is functionally referential, it is necessary to show not only that the call is produced in a particular context, and not only that the call elicits an appropriate response, but also that the

response is specific to the supposed referent of the call (i.e., the call elicits a qualitatively different response than other calls produced in other contexts).

In this study, the authors played back whining calls and anti-predator alarm calls and found that the whining call playbacks resulted in more individuals (both conspecific and heterospecific) approaching the speaker compared to the anti-predator alarm calls. One possible interpretation of these results is that the whining call is a functionally referential call for brood parasites. But an alternative interpretation is that the whining call is simply a high-urgency alarm call, rather than being referentially specific to brood parasites. I don't think it is possible to distinguish between these possibilities with the data provided because the difference in response to whining calls and alarm calls was quantitative rather than qualitative. The birds performed the same behavior (approaching the speaker) in response to both playbacks; they just did it more in response to the whining call playbacks. If brood parasites are a higher-urgency threat than predators (which is consistent with the observation that birds mobbed brood parasite mounts more than predator mounts), this would explain why more whining calls were produced in response to brood parasite mounts. It is worth noting that whining calls were occasionally produced in response to the predator mounts too, which suggests that whining calls are not wholly specific to brood parasites.

To convincingly show that the whining call is functionally referential to brood parasites, I think the authors would need to provide some data showing that birds respond in a qualitatively distinct manner to whining calls and alarm calls. I can think of a few different ways to do this:

- 1) Show that whining calls elicit a unique behavioral response that is appropriate to brood parasites but not to other threats. For example, yellow warblers respond to anti-cowbird seet calls by rushing back to their nest to prevent the cowbird from laying an egg in it. This is qualitatively different from the response to a predator. If whining calls elicit a similar type of behavior, this could be evidence that whining calls are truly specific to brood parasites.
- 2) Conduct playbacks both during and outside of the breeding season. If whining calls are truly specific to brood parasites, we might expect that birds would show a decreased response to whining calls outside of the breeding season when brood parasitism is not an issue. By contrast, the response to anti-predator alarm calls should be constant in and out of the breeding season since predators are a threat year-round.
- 3) Do a prime-probe experiment similar to Zuberbuhler et al. 1999. First "prime" the birds by presenting a mount of a brood parasite or a predator. Then once they have seen the mount, "probe" them by playing back whining calls or anti-predator alarm calls. If the birds mentally associate each call with a particular stimulus, then they should show a decreased response to whining calls after being primed with a brood parasite mount but not after being primed with a predator mount, and they should show a decreased response to alarm calls in the opposite scenario. But if both calls

are just generic warning calls with the whining call being higher urgency, then they should show the same decrease in response to the call playback regardless of which mount they were primed with.

RESPONSE: We thank the reviewer for their kind appraisal of our manuscript and are grateful to them for highlighting this important point. Having read and re-read this comment numerous times, it is clear that we did not: 1) sufficiently describe how past research acts as a foundation for this current study; 2) sufficiently discuss how the results of this (and similar) research calls attention to an unfortunate semantic issue in past research that can muddy the perception of the qualitative distinctness of response towards “whining” vocalizations compared to “mobbing/alarm/terrestrial alarm” vocalizations under some circumstances; and 3) provide sufficiently in-depth analyses to highlight the qualitative (in addition to quantitative) differences in the response host birds make towards “whining” and “mobbing/alarm vocalizations”. To address this comment, we: 1) provide additional information regarding how past research has acted as a foundation for this present study; and 2) provide additional analyses to show that a clear qualitative and quantitative difference in response towards (terrestrial) alarm and whining vocalizations exists.

1) How past research supports this current study

The reviewer raises several important points that we are confident can be addressed by providing a more in-depth explanation of how past research acts as a foundation for this study.

Whining vocalizations are learned signals that refer specifically to cuckoos and are only produced when they present a threat to host nests: Fairy-wrens (Maluridae), and in particular the superb fairy-wren, are the focus of numerous intensive long-term monitoring studies (e.g. 1987-present at the Australian National Botanical Gardens), making them a very well-studied species. Despite numerous studies from various sites investigating behavioral responses by an array of fairy-wren species towards a wide variety of threats (e.g. Magrath et al. 2007 Behav. Ecol. 18: 944-951; Grieg & Pruett-Jones 2009 Anim. Behav. 78: 45-52; Colombelli-Négrel et al. 2010 Behav. 87: 853-870; Teunissen et al. 2021 Curr. Biol. 31: 4120-4126), whining vocalizations remained undocumented in these species until studies specifically began investigating their responses towards cuckoos. The first study to document (and describe) the whining vocalization in any fairy-wren was Payne et al. (1985 Behav. 94: 108-126), who reported it being made by splendid fairy-wrens towards a model cuckoo. Importantly, they presented cuckoo models throughout the nesting cycle (e.g. when the fairy-wrens were building a nest, when there were eggs or nestlings in the nest and when the nestlings had fledged from the nest) and found that

whining vocalizations and extreme aggression (e.g. pecking at the cuckoo's eyes and pulling feathers from its head) were only observed when the cuckoo presented a threat (eggs or chicks in the nest). Whining vocalizations were not produced when cuckoos did not present a threat. The first evidence that superb fairy-wrens produce whining vocalizations was presented in Langmore et al. (2012 *Behav. Ecol.* 23: 798-805), which like Payne et al. (1985 *Behav.* 94: 108-126) presented superb fairy-wren breeding groups with a taxidermied cuckoo. Specifically, Langmore et al. (2012 *Behav. Ecol.* 23: 798-805) showed that superb fairy-wrens that breed at sites that rarely experience cuckoo parasitism (due to these habitats being in human-modified environments that cuckoos typically avoid) rarely produce whining vocalizations or respond with extreme aggression when they see a cuckoo, while those that exist in locations that regularly experience cuckoo parasitism (unmodified environments) produce whining vocalizations at a high rate and respond extremely aggressively towards the cuckoo. As the locations used in this study were within the dispersal distance of fairy-wrens, the authors concluded that the response towards cuckoos is learned rather than innate. Feeney & Langmore (2013 *Biol. Lett.* 9: 20130443) followed this up with a field-adaptation of Curio's classic learning experiment (Curio et al. 1978 *Science* 202: 899-901) and found that cuckoo-naïve fairy-wrens do not respond to the sight of a cuckoo, but do (i.e. they produce whining vocalizations and exhibit extreme aggression) after watching the response of cuckoo-experienced fairy-wrens. Once learned, targeted recognition of and responses towards cuckoos can persist for at least several years (Langmore et al. 2012 *Behav. Ecol.* 23: 798-805). Feeney et al. (2013 *Science* 342: 1506-1508) then specifically examined whether the whining vocalization is functionally referential in the superb fairy-wren, and found: 1) that the whining vocalization is produced in response to a specific referent (cuckoos as opposed to the other threat types that superb fairy-wrens face, Fig. 4C); 2) that hearing this vocalization elicits a qualitatively and quantitatively different and more aggressive response compared to when they hear a mobbing (termed "alarm" in the present manuscript, because while "mobbing/terrestrial alarm/alarm" vocalizations elicit curiosity and vigilance from birds at a distance, they do not actually elicit mobbing, whereas whining vocalizations elicit an extremely rapid and aggressive mobbing response - also see additional analyses, below) or control vocalization (Fig. 4D); and 3) that whining vocalizations are structurally unique compared to other vocalizations within their repertoire (Fig. S3). This study also demonstrated that more mobbing (which whining elicits) correlates with decreased parasitism, highlighting the ecological appropriateness of physical aggression towards cuckoos (discussed in more detail in [2], below).

In addition to facilitating uniquely aggressive mobbing, information about cuckoo presence informs other cuckoo-specific defenses. For instance, Feeney & Langmore (2015 *The Auk* 132: 359-364) showed that the sight of and response towards an adult cuckoo by a superb fairy-wren results in heightened vigilance in the vicinity of their nest in the hours and days

following exposure (like that seen in reed warbler hosts of the common cuckoo in England: Davies et al. 2003 *Anim. Behav.* 65: 285-295). This heightened vigilance following the sight of a cuckoo is the reason extended periods are allowed between model presentation treatments to allow for carryover aggression to diminish; however, this impact does linger, which is likely why published model presentation experiments record occasional instances of whining vocalization production towards other treatments. Additionally, Langmore et al. (2003 *Nature* 422: 157-160; 2009 *Behav. Ecol.* 20: 978-984) showed that information about parasitism risk also influences the likelihood of superb fairy-wrens rejecting a cuckoo chick. While not specifically related to this present study, these studies highlight that hosts of cuckoos exhibit a suite of brood parasite-specific defenses that are calibrated using personal and social information about parasitism risk.

These already published studies together provide evidence that whining vocalizations are learned and functionally referential to cuckoos, and we know of no evidence that contrasts with these conclusions despite diverse research on the topic by numerous research groups over several decades on a well-studied group of birds. In the previous version of our manuscript, we considered that citing these studies would be sufficient for us to extend on them; however, we are grateful to the reviewer for pointing out that a more detailed discussion would be helpful. We have now provided a more comprehensive overview of research on the topic in the introduction to ensure that our logic and assumptions are adequately supported and accessible by readers. We have also provided a more explicit definition of functionally referential vocalizations (see Macedonia & Evans 1993 *Ethology* 93: 177-197).

2) Additional analyses to explicitly demonstrate a qualitative (in addition to quantitative) difference in response to “mobbing/alarm” compared to “whining” vocalizations: Brood parasites do not typically present a threat to adult host birds but do present a threat to their nests. Thus, a highly aggressive response is ecologically appropriate and, according to past literature, an effective defense. For instance, numerous studies have documented how physical mobbing can deter adult brood parasites from parasitizing a nest (reviewed in Gloag et al. 2013 *Anim. Behav.* 86: 1023-1029) and even result in the host killing the adult brood parasite (Molnár 1944 *Aquila* 51: 100-112; Moyer 1980 *Zam. Ornith. Soc. News.* 10: 159; Šulc et al. 2020 *J. Vert. Biol.* 69: 1-6; Zhao et al. 2022 *Ecol. Evol.* 12: e9649). By contrast, predators (e.g. currawongs or hawks) present a mortal threat to adult (host) birds and, thus, physically attacking them likely does not represent an ecologically adaptive response by host adults when these threats also present a threat to the host nest. Consequently, while passerines such as fairy-wrens and scrubwrens make alarm vocalizations from a distance when such a predator is near their nest, possibly to draw their attention away from the nest, we know of no records of them protecting their nests from these types of predators with physical aggression. In the previous version of our

manuscript, we assumed that citing research which demonstrated the whining vocalization as functionally referential towards cuckoos would be sufficient to extend our examination of this vocalization. However, we agree with the reviewer that a more explicit demonstration of the qualitative (in addition to quantitative) difference in the response elicited by “whining” versus “alarm” vocalizations would be helpful for readers of the manuscript.

To demonstrate that hearing whining and alarm vocalizations elicit qualitatively and quantitatively distinct responses in superb fairy-wrens and white-browed scrubwrens that are also ecologically appropriate for the threats they convey information about, we conducted a more in-depth analysis of the behavioral responses we recorded during our 30s playback experiments using the data we previously uploaded to DRYAD. In the previous version of the manuscript, we present data exploring the number of host birds that approached between 0-2m from the speaker (Fig. 2G, H), as our study is primarily interested in exploring an ecologically relevant response towards the whining vocalization, rather than also conducting a more qualitative demonstration of how responses towards the whining and alarm vocalization differ. During these experiments, however, we recorded not only how many birds approached within 0-2m of the speaker, but also how many approached between 2-5m of the speaker. These data were not used, as we did not see them as central to our investigation, but in light of the point the reviewer raises, these data do offer an opportunity to explore whether there was a difference in the number of birds that rapidly (as they are 30s trials) approached the sound source in a manner consistent with the intention to physically attack (0-2m, a distance that also brings with it the risk of being physically attacked) or more cautiously assess (2-5m) the referent to which the sound refers without the receiver putting themselves in danger. Regarding these data, all birds that were recorded to approach 0-2m did so after travelling through the 2-5m distance category, thus presenting a picture of behaviour during the 30s trial period. If whining vocalizations incite an immediate response in birds that are intending to physically attack the referent to which the whining vocalizations refer, we should expect no difference in the number of birds that approach 0-2m versus 2-5m, as it would show that birds rapidly approach the speaker with little concern for their own well-being. Correspondingly, if alarm vocalizations incite a vigilance response rather than an immediate mobbing response, we should expect a difference in the number of birds that approach 0-2m versus 2-5m, as birds approach the speaker close enough to investigate what is being referred to while maintaining a distance sufficient to ensure their own safety. Our data demonstrate exactly this: in response to conspecific whining vocalizations, the same number of superb fairy-wrens and white-browed scrubwrens approached 0-2m as 2-5m ($V = 329$, $P = 0.5913$; $V = 363$, $P = 0.2573$ for superb fairy-wrens and white-browed scrubwrens, respectively), demonstrating that upon hearing whining vocalizations, individuals from both species rapidly (i.e. within 30s) approach very close to the sound source. By contrast, in response to conspecific alarm

vocalizations, we found significant differences in the number of birds approached within 0-2m versus 2-5m ($V = 255$, $P = 0.0003403$; $V = 323$, $P = 0.001214$ for superb fairy-wrens and white-browed scrubwrens, respectively), demonstrating that upon hearing conspecific alarm vocalizations, individuals from both species approach (~half as many as towards the whining vocalizations, see Table 2) but only approach at a distance consistent with them seeking to initially inspect the referent to which the signaler is referring (Table 2).

	Number of conspecifics			
	0-2m	0-5m	V	P
Superb fairy-wren				
whining	2.18 ± 0.21	2.59 ± 0.21	329	0.5913
alarm	1.12 ± 0.17	1.94 ± 0.16	255	0.0003***
control	0.41 ± 0.15	0.59 ± 0.15		
White-browed Scrubwren				
whining	2.06 ± 0.18	2.29 ± 0.14	363	0.2573
alarm	0.71 ± 0.17	1.24 ± 0.16	323	0.0012**
control	0.35 ± 0.12	0.65 ± 0.12		

Table 2. Number of conspecifics that approached between 0-2m and 0-5m of the speaker during 30sec playback experiments at superb fairy-wren and white-browed scrubwren nests ($n = 17$ for all treatments) (mean $\pm$ standard error). Absence of a significant difference and more birds attending whining treatment is consistent with previous findings that the whining vocalization elicits a rapid recruiting response for aggressive mobbing (Feeney et al. 2013 Science 342: 1506-1508), whereas fewer birds approaching and those that did approach doing so to a further distance during alarm vocalizations is consistent with this vocalization eliciting cautious curiosity. Wilcoxon signed rank tests with continuity corrections were conducted in R.

These data also corroborate the data from our model presentation experiment, which shows that the sight of a cuckoo results in aggressive mobbing of the cuckoo (see Fig 2C, D and accompanying analyses), which was a response not recorded in response to the currawong and raptor treatments. Furthermore, when the full dataset (deposited to the DRYAD repository prior to submission of the original manuscript) is examined, it shows that the closest superb fairy-wren and white-browed scrubwren individual spent approximately $\frac{2}{3}$ of the currawong and raptor trials between 2-5m of the model. These data further highlight that the sight of a cuckoo elicits physical aggression, while cautious vigilance is elicited when either species sees a predator that presents a physical threat to the. These qualitatively

(and quantitatively) different responses are ecologically appropriate for the different threats these birds face.

Superb fairy-wren	Time spent by closest individual	
	0-0.5m	2-5m
Cuckoo	237.88 ± 23.89	61.71 ± 23.67
Currawong	14.71 ± 13.06	193.59 ± 27.05
Raptor	45.94 ± 22.81	189.94 ± 26.66
White-browed Scrubwren		
Cuckoo	248.35 ± 20.61	40.12 ± 13.69
Currawong	4.82 ± 3.58	186 ± 26.86
Raptor	1 ± 1	205.65 ± 23.03

Table 3. Time spent between 0-0.5m and 2-5m of the model by the closest individual during the 300s cuckoo, currawong and raptor model presentation treatments (n = 17 for all treatments) (mean ± standard error). Cuckoos do not present a threat to adult birds, while currawongs and raptors are predators of adult birds. Close approach during the cuckoo treatments is consistent with the sight of a cuckoo eliciting a highly aggressive physical mobbing response, whereas the closest bird spending almost no time 0-0.5m but approximately 2/3 of the trial between 2-5m of the currawong and raptor is consistent with the sight of these threats eliciting a cautious vigilance response.

Overall, we are very appreciative of this comment, as it provided us with the opportunity to more completely demonstrate the targeted use and response towards whining vocalizations. In addition to providing a more detailed discussion of past literature in the introduction, we have also incorporated these additional analysis and corresponding discussion of them into the manuscript.

Another concern I have relates to the claim that birds must learn the context in which to produce whining calls. To support this claim, the authors cite two previous publications (Langmore et al. 2012; Feeney and Langmore 2013). They also present an experiment in the supplementary material in which Purple-crowned Fairy-wrens produced fewer whining vocalizations in response to cuckoo mounts in years when brood parasitism was low, suggesting that fairy-wrens who have more experience with cuckoos are more likely to produce whining vocalizations in response to them. However, as far as I can tell, the cited studies and the Purple-crowned Fairy-wren experiment only

suggest that fairy-wrens must learn that cuckoos are a threat. They don't directly address the question of whether whining calls themselves are learned. If whining calls are a generic high-urgency call rather than referring specifically to cuckoos, then it is possible that fairy-wrens innately produce whining calls whenever they feel extremely threatened or upset, and the only thing they need to learn is that cuckoos are a high-level threat.

RESPONSE: Apologies for not being clearer here. Langmore et al. (2012 *Behav. Ecol.* 23: 798-805) present data that shows the response towards cuckoos in superb fairy-wrens is not innate, and therefore must be learned. Feeney & Langmore (2013 *Biol. Lett.* 9: 20130443) extend on this by explicitly demonstrating that cuckoo-naïve superb fairy-wren individuals learn to respond aggressively towards a cuckoo, including producing whining vocalizations (we know of no evidence that they produce them in any other context) and physically attacking the model, by watching cuckoo-experienced individuals respond to a cuckoo model. While several studies have presented indirect evidence of hosts socially learning to respond towards brood parasites (reviewed in Campobello et al. 2017 In: M. Soler: *Avian brood parasitism*. Springer International Publishing), Feeney & Langmore (2013 *Biol. Lett.* 9: 20130443) is the only study that we know of that has directly demonstrated a socially learned response towards cuckoos in individuals of any host species by using a methodology that resembles Curio's classic social learning experiment (Curio et al. 1978 *Science* 202: 899-901). The purple-crowned fairy-wren data we present does resemble that presented in Langmore et al. (2012 *Behav. Ecol.* 23: 798-805); however, it is intended to be considered in the context of this and Feeney & Langmore (2013 *Biol. Lett.* 9: 20130443), which is consistent with a learned response. To address this comment, we have more explicitly outlined the findings of previous research and how they act as a foundation for this current study throughout the revised text.

Other comments

- For your analysis of macroecological correlates of whining call production, how did you definitively determine that a species does NOT produce whining calls? Did you have comprehensive data on the vocal repertoires for all these species? Is it possible that some of them do produce whining calls but it simply hasn't been documented yet?

RESPONSE: Vocal responses towards brood parasites are rarely reported, poorly curated and lack a consistent nomenclature, which is potentially why whining vocalizations have remained unexplored for so long, despite our finding that they are a common response in a wide array of host species towards various brood parasite species, including from several brood parasite lineages across several continents. Prior to this study, whining vocalizations towards brood parasitic cuckoos had only been documented in the literature in the splendid fairy-wren (Payne et al. 1985 *Behav.* 94: 108-126) and superb fairy-wren (Langmore et al.

2012 Behav. Ecol. 23: 798-805; Feeney & Langmore 2013 Biol. Lett. 9: 20130443; Feeney et al. 2013 Science 342: 1506-1508), with strikingly similar vocalizations towards brood parasites and corresponding behavioral responses also documented in large-billed gerygones (Noh et al. 2021 Anim. Behav. 174: 51-61) and several Old-World leaf warblers (Wheatcroft & Price 2015 Behav. Ecol. 26: 83-90). In this study, we identified species that do produce whining type vocalizations by obtaining recordings and analyzing the features of vocalizations made towards brood parasites versus those made towards predators (see the *Vocalization dataset* and *Acoustic analysis* sections of our materials and methods for more details). This analysis determined that 21 of 26 species produce “whining” type vocalizations, while 5 do not. For our macroecological analysis, we complemented this dataset with the results of a literature analysis (see the *Ecological correlates of whining vocalization production* section of our materials and methods). The studies included in the literature analysis were not accompanied with recordings, and thus we were only able to investigate whether they noted unique vocalizations towards predators versus brood parasites (3 of 32 species from 39 studies reported unique vocalizations). If no unique vocalization was recorded, we coded them as no evidence of them producing whining vocalizations, as all evidence suggests that this is a unique response towards brood parasites (outlined in detail in response to your first comment) and thus the lack of a unique response constitutes current evidence of no whining vocalization. It is possible that some of these (and other) species will be found to produce whining (or other unique) vocalizations in the future following more detailed studies; however, we can only work with data that is currently available, and we are confident that our approach is appropriate and conservative. Regardless, this comment highlights that more detail regarding our logic was warranted, and we have now added more information in the methods and to the caption of Fig 1, accordingly.

- Line 243: You say that “more individuals from other species responded”, but it is not clear from the way this sentence is worded what they responded to. Which playback stimulus elicited a greater response by other (non-yellow warbler) species?

RESPONSE: Thank you for pointing this out. We have added “to superb fairy-wren whining vocalizations” to make this clearer.

- Lines 335-336: You say that out of 24 species you had 17 with 2 recordings per context and 8 with 1 recording for one or both contexts. But $17 + 8 = 25$, not 24. What explains this discrepancy?

RESPONSE: Thank you for noticing this. The discrepancy was a simple mistake due to us including the drongo in the list of species with recordings, but not the total number. We provide an explanation as to why this species was excluded in the *Vocalization dataset* section of the materials and methods (in short: it is a vocal mimic with a very diverse array

of vocalizations, which makes classifying vocalizations very difficult). We also note that in the updated version of the manuscript we have included two new species in our analyses (variegated fairy-wren, *Malurus lamberti* and fan-tailed gerygone, *Gerygone flavolateralis*) as data for these species recently became available to us (collaborators who provided the recordings have also been added as coauthors). Accordingly, we have updated/corrected all the relevant sample sizes.

- Lines 358-359: What features were included in this LDA (the one with 85% accuracy)? Was it the first 10 principal components? Please clarify this in the text.

RESPONSE: The LDA includes the first 10 principal components. We have updated the text to clarify.

- Line 524: How do you know that 30 minutes was enough time to wait in between trials? It doesn't seem like much time

RESPONSE: This interval has been established by similar audio playback experiments, including in our own previous research (Feeney et al. 2013 Science 342: 1506-1508). Thank you for pointing out that we had not cited literature in support of this. We have added the reference to this sentence, accordingly.

Reviewer #2 (Remarks on code availability):

When I tried to access the data and code through the link provided to Dryad, the link did not work. Also, I do not have the requisite expertise to assess the phylogenetic methods used.

RESPONSE: Apologies, the data had been uploaded to DRYAD prior to submission, but we did not have the correct settings applied to enable reviewers to access the data/code. As soon as this was brought to our attention by the Editor we changed the settings.

Reviewer #3 (Remarks to the Author):

Overall assessment:

Using a variety of methods, the MS shows that whining calls produced by hosts of brood parasites are similar acoustically and functionally across species. The study combines network analysis with playbacks experiments and attempts to come up with a unified theory about these signals as functionally referential. I am not sure I agree with the argument about referential signaling and its relations to evolution of human spoken language, I think these claims are somewhat speculative, but I find the results interesting nonetheless. Overall, the study is well conducted and well

presented, but several issues in the statistical analysis worries me a bit and requires careful attention before this MS can be published.

Major concerns:

1. A major argument in the MS is that species that occur in areas where the underlying brood parasite–host interaction network shows low interactions per species, rarely produce whining vocalizations. But I thought that, by design, network interactions are computed, at least in part, based on vocal interactions, is this correct? If so, this entire claim feels a bit tautological. Authors should clarify how is host interaction network computed, and how they validate independence.

RESPONSE: Thank you for pointing out that we could have been clearer here. To clarify: network interactions are computed entirely independently of the acoustic data. Briefly, brood parasite—host interaction networks were generated using: 1) documented host status (i.e. evidence of brood parasite eggs or chicks in the nest of a heterospecific), and 2) evidence that a brood parasite and a host exist in sympatry (following Kennerley et al. 2022 *Ecol. Lett.* 25: 1889-1904). Vocalization data were collected independently by researchers or from sound recordists who deposited their recordings in online wildlife media databases, such as xeno-canto (<https://xeno-canto.org>). To examine whether ecological variables are associated with whining vocalization production, we investigated whether there was a relationship between whether a species' produces a whining call or not, and its cooperative breeding status (a measure of the role of intraspecific interactions, obtained from Cockburn 2006 *Proc. R. Soc. B* 273: 1375-1383) and the median linkage density across the species' range (a measure of the role of interspecific interactions: Bersier et al., 2002 *Ecology* 83: 2394-2407; here applied to brood parasite-host networks, following Kennerley et al. 2022). The manuscript has been modified to comprehensively explain what brood parasite—host network interactions show, how they were computed and why they are independent of vocal interactions.

2. Certain claims seem unsupported by data presented. For example, the title of figure 2 is “Whining vocalizations are functionally referential in multiple species”, but in fact, all the figure shows is response to playbacks in two species (one of which has been studied earlier). The rest of the argument is based on panel K, which only show a list of neighboring species that interacted as well, but with no quantification. In fact, I could not understand much from panel K. Overall, I think that the MS can improve if authors should show some level of modesty in their claims.

RESPONSE: We toned down the language used to improve its arguments and conclusions, as suggested. Regarding the title of figure 2, we agree and have changed “multiple species” to “superb fairy-wrens and white-browed scrubwrens” to be more accurate. Apologies for any confusion, but in addition to presenting the results of playback experiments (G-J), this

figure also presents the results of model presentation experiments (A-F) as well as the results of point counts at host nests (K). Importantly, the combination of model presentation and audio playback experiments show that “whining” vocalizations are produced towards cuckoos and that “whining” vocalizations elicit a distinct response relative to that elicited by other vocalizations. Extending on this, panel K compares the species that are found near host nests with those that attended (i.e. came within 5m of the nest) cuckoo model presentation experiments, which allowed us to examine commonalities in the species that responded to “whining” vocalizations. The results here show that passerines, which tend to be hosts of most avian brood parasitic lineages, were more likely to attend than non-passerines. We agree that more information is helpful to highlight what K shows and have changed the caption accordingly.

3. There is some confusion in my mind about many of the statistical claims made in the MS. The authors have a bad habit of presenting p values and chi2 without presenting the means for each group. Furthermore, colormaps are often not explained and are not, in any direct manner related to raw measures that can be easily understood. See specific comments below.

RESPONSE: Ordinarily we would have provided these numbers in the main text; however, we did not as it is impractical for our datasets. For example, providing means and standard errors for Figure 2 constitutes 42 means (plus another 42 standard errors) and another 32 means (plus another 32 standard errors) for Figure 3. We agree that this is a valuable point, so we have now provided the most relevant means in the text and additionally provided supplementary tables that include means $\pm$ standard errors for all experimental results. Additionally, regarding the heat map in Figure 1, we have provided more explanation in the figure legend to help ensure that it is easier to understand.

Specific comments:

1. Acoustic analysis:

This statement confuses me:

“We analyzed the acoustic features that distinguish whining vocalizations from alarm vocalizations and found that the linear discriminant function loaded most strongly with PC3, which was negatively associated with frequency bandwidth and positively associated with spectral flatness.”

If this is a new measure, perhaps it would be better to develop it on a training set and testing it on a different experimental data set? Otherwise, it is sort of ad hoc, and statistical significance of this acoustic distance is not robust. Please explain.

RESPONSE: This is a good point about the lack of validation and testing sets. To address this comment, we have now re-structured the analyses. The brood parasite-associated vocalizations of the superb fairy-wren (Feeney et al. 2013 Science 342: 1506-1508) and two *Phylloscopus* species (Wheatcroft & Price 2015 Behav. Ecol. 26: 83-90) have previously been shown to produce brood parasite-specific responses in receivers. They have also been shown to be functionally referential (Feeney et al. 2013 Science 342: 1506-1508) and be acquired through social learning (Langmore et al. 2012 Behav. Ecol. 23: 798-805; Feeney & Langmore 2013 Biol. Lett. 9: 20130443) in the superb fairy-wren, and that their whining vocalizations have clear acoustic similarities to those seen in the two *Phylloscopus* warblers (Wheatcroft and Price 2015 Behav. Ecol. 26: 83-90). Thus, we now first perform a permutation-based linear discriminant analysis, in which we train a linear discriminant function to classify calls to contexts using a subset of brood-parasite and terrestrial alarm calls from these three species. We validate the function on the remaining calls from these three species and show that it almost perfectly classifies calls to the correct context. We then test a linear discriminant function on the brood-parasite and terrestrial alarm calls from the remaining species. As for the original analyses, this linear discriminant function loaded most strongly with PC3. We then use the linear discriminant function to assign linear discriminant scores to each call to compare call types within species. We outline this procedure and its concordance with our original findings in the revised methods section and include the code in our revised submission.

This one through me off:

“Using the function `lda MASS v7.3-6050`, we generated a single linear discriminant function, which correctly assigned 85% of vocalizations to context”

what is this function exactly? It is not enough to say you used a certain toolbox, you need to document the exact procedure here

RESPONSE: We now provide the exact procedure. We both explain our procedure in more detail in the methods and provide the R code used to generate the function together with the submission, which should now be available to all reviewers (there may have been some issues with accessing the R code in the initial submission for which we apologize).

Same here:

“A species was considered to produce distinctive vocalizations in parasite and predator contexts when the linear discriminant scores of their vocalizations differed in post-hoc tests run using `emmeans v1.8.752`.”

Seriously, is this something a reader should be able to grasp? Here again, I don't understand the procedure, it feels as if the authors are not motivated to help the reader to understand what they did but expect reader to trust that the methods are okay. Most concerning is the binary decision

made here: why binarizing the data rather than using an acoustic distance metric derived from PC3?

RESPONSE: We appreciate you pointing out that we could have been clearer here. We have clarified the methods now in the text to avoid confusion (and note that the R code used to generate these results is included with the submission). Briefly, we constructed a regression model to explain variation in linear discriminant scores through an interaction between calling context and species. Subsequently, we used posthoc tests to determine which species have significantly different linear discriminant scores across contexts. We discuss the reviewer's question about using PC3 vs the binary classification in our following reply.

Finally, reading this statement:

“The linear discriminant function loaded most strongly with PC3 (Data S6), which, considered alone, correctly classified 91% of vocalizations”

This is, in fact, higher than your discriminant function! I wonder if the MS can improve if authors just use PC3 as a metric – as is – no categorization, etc. A color map of PC3 would be much easier to interpret in my view. As it is, in Fig 1 (The global distribution of distinct or non-distinct vocalizations towards brood parasite), I cannot understand this color map. What is exactly corresponding to these colors? I would find it much easier to interpret if these color map values are simply the mean PC3 values in each location.

RESPONSE: We apologize for not being clearer regarding our approach in the previous version of the manuscript.

First, to clarify: the discriminant function classified vocalizations more accurately than PC3 alone. In the original submission, we (1) constructed a linear discriminant function using all 10 PCs to classify vocalizations using all 24 species (85% accuracy), (2) determined which species produce brood parasite-associated calls that get classified together (i.e. ‘whining vocalizations’), (3) constructed a second linear discriminant using PC3 alone on this subset of 20 species (91% accuracy). Based on either the dataset of all 24 species or the subset of 20 species, the linear discriminant performed better than PC3 alone.

However, following your first specific comment, we have now taken an entirely new approach to this analysis (see response above). Importantly, despite this new approach our results remain the same. Regarding the colour map, blue refers to species that our analyses have determined to have whining vocalizations towards their brood parasites. In addition to taking a new approach to our analysis, we have provided more information regarding the methods for these analyses as well as the caption for Figure 1.

Similarly, here:

“Of the 24 host species analyzed, 22 species produced distinct vocalizations towards brood parasites”

This is a highly derived measure. It is sufficient to refer to the PC3 values in figure 1 (much more convincing in my view).

RESPONSE: As discussed in response to your previous comment, the linear discriminant function is more accurate at classifying vocalizations than PC3 alone. Following your first specific comment, however, we have now taken an entirely new approach to this analysis (see response above), which we feel is much more robust while delivering the same result.

Some points than need further clarification:

“If within-species cooperation is a primary benefit of whining vocalization production, species that produce them should be more socially cooperative^{26,37}. Alternatively, if between-species cooperation is a primary benefit of whining vocalization production, species that produce them should reside in areas that have complex brood parasite-host interaction networks”

Why cooperative breeding helps here exactly? I guess in general, cooperative breeding allows more guarding, but even just having a colony of close by nests could help here. So I feel that within and across species networks are not easily distinguishable here. Everyone gains from the warning whining, and the more proximal the nests, the more they gain, isn't it right?

RESPONSE: A series of studies (including, but not limited to, those cited in the above quote from the manuscript) have demonstrated a reciprocal relationship between brood parasitism and cooperative breeding in birds, whereby cooperative groups are better able to defend their nests against brood parasitism, but also offer a larger reward to brood parasites as cooperative groups are often more successful breeders than their pair breeding counterparts. Given this, and that (as stated by the reviewer) information about whining providing general information to all nearby species, our logic is to test between these alternate hypotheses. We appreciate the reviewer for highlighting that more text would be useful to better explain our logic regarding cooperative breeding, so in addition to providing more references in support of this point, we have amended the sentence to: “If within-species cooperation is a primary benefit of whining vocalization production, due to larger groups being better able to defend their nests against brood parasitism (Feeney et al. 2013 *Science* 342: 1506-1508; Canestrari et al. 2009 *Anim. Behav.* 77: 1337-1344; Trnka & Prokop 2011 *J. Av. Biol.* 42: 192-195; Brown & Lawes 2007 *Ibis* 149: 321-327; Wells & Barker 2017 *Anim. Behav.* 133: 47-56), species that produce them should be more socially cooperative.”

“complexity of underlying brood parasite–host network”

Please explain the proposed mechanism here, complexity is a complex term...

RESPONSE: Brood parasite—host networks refer to relationships between brood parasite and host species occurring in a given area. Networks can be broadly categorized as one of four types: one-to-one (one brood parasite and one host), one-to-many, many-to-one and many-to-many; however, one can also calculate a more precise measurement of a given network to compare networks on a continuous scale. In Kennerley et al. (2022 *Ecol. Lett.* 25: 1889-1904), we used linkage density (i.e. the number of interactions divided by the total number of species) to quantify network complexity, which is a commonly used descriptor for network complexity (Bersier et al., 2002 *Ecology* 83: 2394-2407). We have added additional wording in the methods explaining this.

“opportunity for between-species cooperation”

How is this quantified exactly?

RESPONSE: When host species occur in sympatry there is the opportunity for interspecific communication of the threat of a brood parasite (e.g. producing brood parasite-specific whining vocalizations), which could provide a fitness benefit as information about parasitism risk is known to be used to calibrate defenses in an array of host species (e.g. Davies & Brooke 1988 *Anim. Behav.* 36: 262-284; Davies et al. 2003 *Anim. Behav.* 65: 285-295; Welbergen & Davies 2009 *Curr. Biol.* 19: 235-240; Langmore et al. 2009 *Behav. Ecol.* 20: 978-984; Feeney & Langmore 2015 *The Auk* 132: 359-364). To quantify the opportunity of interspecific cooperation, we use linkage density of brood parasite–host networks, our measure of network complexity, as a proxy (described in response to previous comment). This is because networks which have the greatest connectivity and complexity should be those that are most likely to promote the evolution of interspecific communication about the threat of parasitism. We have further clarified in the methods that we use linkage density of brood parasite–host networks as our measure of the opportunity for interspecific cooperation.

“Fig 2E: Distribution of species that attended cuckoo model presentation experiments or were detected in the vicinity of host nests. Blue shading indicates whining”

There is no blue shading in K!

RESPONSE: The text sentence reads: “Blue shading indicates whining vocalizations produced in response to the sight of a model cuckoo (A-B), or behavioral responses towards whining (G-J) playbacks.”.

Finally:

The statistical tests should all show mean values for each group, not only chi sq. This should be consistently presented in the entire MS

RESPONSE: This was answered in response to comment 1, above. Briefly, we have now provided the most relevant means in the main text (all is impractical as it would constitute 74 means [82 including new analyses] and the same number of standard errors). We have also provided supplementary tables that include summaries of our experimental data.

Reviewer #4 (Remarks to the Author):

Review of "Learned use of an innate sound-meaning association in birds"

The authors present a series of experiments to test whether whining vocalisations associated with a brood parasite in one songbird system evoke a response to that call in other songbird systems in different geographical areas - even if the last common ancestor was 53 million years ago. The paper fills an important gap in our conceptual framework of the phylogenetic history of learned sound associations. The authors show that some innate vocalisations elicit a similar response in other (naive) bird species in allopatry. The authors conclude that their findings are parsimonious with the view that learned sound symbolism derives from innate origins.

The main weakness of the manuscript in its current form is the treatment of alarm calls. In line 106, the authors write “alarm vocalisations (that) do not typically refer to a specific referent”. But Cheyney and Seyfarth have shown referential alarm calls in vervet monkeys, and songbirds have aerial and terrestrial alarm calls that differ in frequency characteristics. There are many (review) papers on referential alarm calling. Can you clarify what you mean, as you elicited alarm calling with a model snake, falcon, and currawong? Perhaps the model birds were on the ground, which is known to elicit terrestrial rather than aerial alarm calls.

RESPONSE: This is a great point, which highlights the published literature’s inconsistent use of terminology for alarm vocalizations across study species that appears to have contributed to some confusion with this manuscript. We appreciate the reviewer pointing it out and we agree that we could have been clearer here. To clarify, we (as the reviewer suggests) do refer to what is sometimes referred to as terrestrial alarm vocalizations; however, we note that in research into anti-predator vocalizations in fairy-wrens (which predates Cheyney & Seyfarth’s fantastic work), they are simply referred to as “alarm” calls/vocalizations and are made generally towards terrestrial threats (see Rowley 1965 *Emu* 64: 251-297 through Teunissen et al. 2020 *Curr. Biol.* 31: 4120-4126). Also, as correctly speculated by the reviewer (and outlined in the methods) all predators, including those that can elicit aerial alarm vocalizations in the correct context, were in a perched posture and placed on the ground/perch during field experiments, and thus elicit “terrestrial” alarm vocalizations. We have now made it clear in the manuscript that “alarm” vocalizations refer to terrestrial alarm vocalizations.

Abstract

I found the abstract rather choppy. In line 3 you refer to "similar vocalisations", but one only understands the meaning after reading the manuscript. In line 5 you write "this vocalisation globally", which is again confusing. I think you mean over large geographical areas. In line 6 you write "this vocalisation is functionally referential", but you mean "this vocalisation type" (referring to the whining call type). Your repeated reference to "this vocalisation" makes it seem as if there is one vocalisation, but you mean the whining call type with similar frequency characteristics across geographical areas. I know space is limited -- but the abstract was hard work. Finally, in line 14, you write "is learned for its production", and again, if possible, clarify "sound production". The last sentence would benefit from adding "of sound production" at the end.

RESPONSE: We appreciate this comment. Substantial edits have now been made to the abstract to better explain what we did, why we did the things we did, and what our results show. Yes, space is limited, but we had room to expand and now use the allowed abstract word limit for this journal.

Line 52: the term innate vocal reactions appears for the first time, and then throughout the manuscript. I think I understand what you mean, but what is a little difficult to quickly grasp is the fact that the "reactions" you refer to are, as I understand it, reactions to internal states or external triggers that result in innate vocal production. So innate vocal reaction is equivalent to innate vocal production (as the consequence of a reaction to some cue or state). Is there any sense to make this point here?

RESPONSE: We have added more explanation here. We have now included a more explicit example to illustrate what innate vocal reactions are: "For instance, screams in humans and functionally similar distress vocalizations in other species are innate vocal reactions...".

Line 74: usually we think of selection on honest signalling, not on honest responding. Maybe delete "honest" before "responding"? The predictability is important for the statistical probability that may underpin forms of associative learning.

RESPONSE: Changed.

Line 87: might be worth noting somewhere that you are referring to vocal production learning systems. The complication with the terminology is that the so-called "vocal non-production learning systems" can have associative sound learning.

RESPONSE: This is a good point, thank you. This is now noted in the second paragraph of the introduction.

Line 92: you write “but their response to it is innate”. Is the vocal production also innate?

R: Production of the vocalization and an aggressive physical mobbing response towards brood parasites is learned (see Feeney & Langmore 2013 Biol. Lett. 9: 20130443). It is clear from this, and other comments, that more discussion of past research was needed. The text has been substantially modified to highlight the results of past research and how they provide a platform for this current study.

Line 96: replace word “using” with “producing” a similar sound solution

RESPONSE: Agreed and changed.

Lines 104-106: This issue of alarm calls is the biggest weakness of the manuscript. Clearly, birds and mammals have referential alarm calls. In songbirds, aerial alarm calls tend to be whistles, and terrestrial alarm calls tend to be broad frequency bandwidth calls, where urgency is also communicated by call rate.

R: Yes, while we are using phrasing that is convention with this study system/species, we see now that more information regarding our decision for using the wording “alarm” vocalizations in the previous version of the manuscript would have been helpful. Briefly (as it is also discussed above in response to a previous comment), we use “alarm” as it is established convention to use this terminology when referring to terrestrial anti-predator vocalizations in fairy-wrens (e.g. see Rowley 1965 Emu 64: 251-297 through Teunissen et al. 2021 Curr. Biol. 31: 4120-4126). As clarified above (and in the text) what we refer to as “alarm” vocalizations are often referred to as “terrestrial alarm calls” in other systems. Like in other systems, these calls refer generally to terrestrial threats in fairy-wrens (e.g. Rowley 1965 Emu 64: 251-297; Teunissen et al. 2021 Curr. Biol. 31: 4120-4126). Alarm calls differ from whining calls, which are specifically used towards cuckoos, as shown in this paper and other past research (e.g. Feeney et al. 2013 Science 342: 1506-1508 provides an explicit examination of this in the superb fairy-wren; however, it has also been described, again only towards cuckoos when they present a threat, in the splendid fairy-wren, Payne et al. 1985 Behav. 94: 108-126). We have now amended the manuscript to indicate that we specifically refer to those towards terrestrial predators.

Line 116: which predator types and hence which predator-associated alarm call types?

RESPONSE: Manuscript has been updated to indicate terrestrial predators.

Line 126 and 133: which predator-associated alarm call types?

R: Following your previous (very helpful) comment above, we have clarified that the predator-associated alarm vocalizations that we use as a comparison in this study refer to those often termed terrestrial alarm vocalizations. After describing more fully what we are using (terrestrial alarm vocalizations), we indicate that from there on they will be referred to simply as “alarm” vocalizations, which is what they are referred to here. As they are not the focus of this study, just an object for comparison, we believe that referring to them in this way after appropriately defining them should be sufficient given that there is existing precedent for this terminology (e.g. Rowley 1965 *Emu* 64: 251-297 through Teunissen et al. 2021 *Curr. Biol.* 31: 4120-4126).

Line 138: innate danger? Perhaps rephrase as innate vocal response to a threat.

RESPONSE: We agree and have changed the text to “innate response”.

Figure 1: there is no ‘general alarm call’ that I am aware of. There are terrestrial and aerial alarm calls, for example.

RESPONSE: We have amended this to “terrestrial alarm vocalization”.

Line 141: insert word ‘host’ before defenses

RESPONSE: Done.

Line 161: why global? I think you mean “with the geographical distance of species”

RESPONSE: We originally used the word global here as we included data from all continents except Antarctica (because no obligate brood parasite-host interactions occur there). Nonetheless, we agree with the point here and have changed the word to “geographic”.

Line 170: replace “globally” with “at larger geographical scales”, or something like that.

RESPONSE: While we agree with the previous change, we have not made this change here as our analysis is looking at the global distribution of traits and includes all areas where these interactions occur.

Lines 188 and 194 etc – the general alarm call issue.

RESPONSE: We have now added “terrestrial” in front of both instances.

Line 242: insert “to whining calls” after “other species responded”

RESPONSE: Done.

Line 321: when you refer to the median number of vocalizations per species – do you mean host species or predator species?

RESPONSE: Thank you for pointing out that we could have been clearer here. The sentence now reads: “recordings of vocalizations in response to brood parasites... and in response to predators...”

Lawson et al 2023 in *Behavioural Ecology and Sociobiology* showed that Galapagos Yellow Warblers do not produce the anti-brood parasitic seet call, nor do they respond to an experimental broadcast of the North American Yellow Warbler's seet call. The authors conclude that this supports the view that the association between the seet call and a brood parasite is learned, as you write. How does this island-endemic response pattern fit with your findings? This might be interesting to consider, as the Galapagos Yellow Warblers are thought to have diverged recently, within the last 300,000 years. Or perhaps the whining call is a special class of calls?

RESPONSE: We appreciate this comment. In short: yes, the results of Lawson et al. 2023 *Behav. Ecol. Socio.* fit perfectly with the finding of this, and other similar studies. For instance, Langmore et al. (2012 *Behav. Ecol. Socio.* 23: 798-805) found that superb fairy-wrens in locations that rarely experience brood parasitism (i.e. within human modified environments, which cuckoos typically avoid) rarely produce “whining” vocalizations or respond aggressively towards cuckoos, while those in nearby populations that experience high brood parasitism rates produce “whining” vocalizations and are extremely aggressive towards them. Furthermore, Feeney & Langmore (2013 *Biol. Lett.* 23: 20130443) demonstrated that individual fairy-wrens that have never seen a cuckoo do not respond to them, but do (including producing “whining” vocalizations and being extremely aggressive towards them) after watching cuckoo-experienced fairy-wrens respond towards a model cuckoo. Basically, all research into brood parasite-specific vocalizations indicates that the production of them towards the correct referent is learned. What makes our study here stand apart is that despite the production of “whining” vocalizations in the correct context being learned, the response towards hearing it is innate. To our knowledge, the whining vocalization represents the first vocalization known to contain learned and innate features in this way, which is why it is of interest to linguists as it resembles a conceptual “missing link” between learned-arbitrary (i.e. words) and fixed-automatic (i.e. typical animal

vocalizations) vocalizations. In response to this, and other similar comments by the other reviewers, we have modified the text to better highlight how our results build on the findings of past research, and we have also cited Lawson et al. (2023 Behav. Ecol. Socio. 23: 798-805) in the introduction when discussing that referential alarm vocalizations towards cuckoos are learned in all systems that have been studied to date.

Can you provide spectrograms of the various call types across geographical areas, including examples of the alarm calls for the experiments?

RESPONSE: We have now provided spectrograms of both call types for all 26 species in Supplementary Figure S1.

Reviewer #5 (Remarks to the Author):

This paper examines the occurrence of a brood parasite specific alarm call given by hosts of brood parasites (the whine call) and frames the study as important in terms of the evolution of human language. The authors first point out the features of Darwin's idea about language evolution and then point out the possible connections with the alarm call of cuckoo hosts. They then explore several aspects of the whining call. First, they show that the whine calls are taxonomically widespread, that these calls given by various hosts of brood parasites are all similar in structure to each other but differ from alarm calls given to predators. Second, they do a comparative (macroevolutionary) study that shows that the whining call is predicted not by cooperative breeding (and hence communication within groups) but by the richness and complexity of the brood parasitic community a give host lives in (which they speculate indicates that calls function in heterospecific communication among hosts). Third, playback experiments combined with model presentations provided evidence that a second species (scrubwren) uses whining calls in the context of brood parasitism (previously shown for fairy-wren). Finally, playbacks showed that species respond appropriately to whine calls even from taxa they are allopatric with

In terms of positives, this paper explores a very interesting phenomenon, that hosts of brood parasites appear to use an alarm call that is specific to brood parasites and that this call may have evolved early in the songbirds or independently in various taxa. There is some evidence for this claim, but also some clarification or additional information is needed for this to be fully convincing. I found the paper a bit of struggle to read and to see how the parts fit together. Some of this could be because the initial framing of the study adds an unnecessary complication in terms of having the authors try to tick the boxes for that framing rather than focusing entirely on making a convincing case for the use of specific alarm calls aimed at cuckoos. To be fair to the authors and make sure that I was not being overly crabby about the paper I shared it with a colleague (name provided to the editor) who is an expert in alarm calls and mobbing behavior in birds. That colleague had a very similar response, both overall and to specific points I outline below. I do not

believe the paper is publishable in its current form and it needs to be considerably reframed, and certain aspects need to be made clearer with additional details.

1. In my opinion, the framing of the paper as presenting an example that supports Darwin's musings about language evolution does not work. This framing made the study harder to follow because the reader is focused on ensuring that the assumptions that would support Darwin are met, rather than focusing on the topic of a specialized brood parasite alarm call itself, which is interesting in its own right. This framing makes the paper a bit hard to follow in places and gets in the way of clarity. In addition, the alarm call seems like a bad fit with the human language framing because human words are learned, whereas the alarm calls are innate. The fact that these calls 'sit between the extremes of innate and learned' (line 90) is irrelevant. The fact that the context is learned, but not the sounds themselves, means the connection to human language does not work.

RESPONSE: As the product of a collaboration between behavioral ecologists and evolutionary linguists, we have been mindful of writing the manuscript in a way that is accessible to readers from both disciplines, and we are happy to see that the other reviewers praised the quality of the writing and/or way the rationale of how the paper was laid out. We appreciate this comment, as well as those from the other reviewers that highlight how we could have been clearer in our logic and we have made numerous comments to address this throughout. Importantly, all published literature on referential alarm vocalizations towards brood parasites indicates that they are learned (e.g. Langmore et al. 2012 *Behav. Ecol.* 23: 798-805; Feeney & Langmore 2013 *Biol. Lett.* 9: 20130443; Lawson et al. 2023 *Behav. Ecol. Socio.* 77: 99), and we know of no published evidence that suggests otherwise. Given that whining vocalizations have been shown previously (in addition to the results of this study) to be functionally referential towards cuckoos (Feeney et al. 2013 *Science* 342: 1506-1508) and that they are only produced towards cuckoos during the breeding cycle when they present a threat to host nest contents (Payne et al. 1985 *Behav.* 94: 108-126; Langmore et al. 2012 *Behav. Ecol.* 23: 798-805), these studies highlight that, as far as experiments can disentangle, the vocalization and the context in which it is produced is learned. Regarding the relevance of our work to linguistics, we note that the senior author of our paper is a linguist and that we have thanked or otherwise acknowledged several other linguists who provided comments on our work prior to us submitting it. Finally, as Darwin was the first (that we know of) to suggest an evolutionary pathway between animal vocalizations and human language potentially exists, we consider it appropriate to credit him accordingly. That said, the reviewer's comments do raise the important point that Darwin did not set out specific hypotheses to test this idea. Correspondingly, we have amended the text at the end of the second paragraph of the main text to reflect this and de-emphasize Darwin.

2. A concern for me was that I was not fully able to consider the results of this study in the context of the literature on mobbing vocalizations and to be convinced that what is shown here is clearly distinct. There are a few concerns. First, the way the hosts respond to cuckoos seem like a mobbing reaction and the vocalizations are somewhat similar in structure to mobbing calls. Despite this similarity, the term mobbing is never used in the paper and possible connections to predator mobbing are not discussed. Second, it is well known that many songbirds (at least temperate zone birds) have two distinct types of alarm calls, a pure tone hard to locate seet call that signifies an immediate threat like a flying raptor, and a more broadband with harmonic structure call that is used in mobbing, is easily to locate for recruiting other mobbers. The authors lump all predator alarm calls together and it is hard to know if how much of the difference between whines and alarm calls is because they have used mainly seet type calls. For example at line 104 "To test this, we collated and analyzed the acoustic structure of vocalizations produced by hosts towards brood parasites and predators, as predators also present a threat but are typically met with alarm vocalizations that do not typically refer to a specific referent (hereafter 'alarm' vocalizations)". I think using mobbing type calls and not seet type calls would be more appropriate for this comparison. Note that if seet calls were compared to mobbing calls in many songbirds they would show up as strikingly different, more so that the differences shown in the current paper where it was the third level of the PC (PC3) where separation was shown. Perhaps the distinction I am making here (mobbing vs seet calls) does not apply to some of the taxa investigated here but these details should at least be considered and discussed.

RESPONSE: These are two important points, with point one being raised (in one form or another) by several reviewers.

Regarding the first point (vocalization nomenclature): As most comprehensively answered in response to reviewer 2 (but also see our responses to reviewer 4), there are terminological issues regarding what word(s) should be used for mobbing (R5) / terrestrial alarm (R4) / alarm (R2) vocalizations. Despite us using established naming conventions for our study species (e.g. see Rowley 1965 *Emu* 64: 251-297, which used "alarm" vocalization to refer to "terrestrial" threats, a terminology that has been consistently used since, see Teunissen et al. 2021 *Curr. Biol.* 31: 4120-4126) it is evident that we should have been clearer regarding our logic. As these vocalizations closely resemble what is often termed "terrestrial alarm vocalizations", we have added this to our initial description in the manuscript, following reviewer 4's suggestion. We have also intentionally not used the word mobbing vocalizations, as our analyses indicate that (terrestrial) alarm vocalizations and whining-type vocalizations are structurally distinct (also see Feeney et al. 2013 *Science* 342: 1506-1508 for analyses regarding this in superb fairy-wrens in a different location, almost 2000 km south of the study site used in the present manuscript). Additionally, and importantly, our analyses also demonstrate that the (terrestrial) alarm vocalizations do not promote a mobbing response (despite this vocalization occasionally being referred to as a

“mobbing” vocalization), rather a vigilance or inspection response (see new analyses added in response to reviewer 2’s comments), while the whining-type vocalization promotes an immediate and very aggressive mobbing response that is both qualitatively and quantitatively distinct from alarm vocalizations, while also being an ecologically appropriate response towards the threat that brood parasites present.

Regarding the second point (vocalization types): we agree that we could have been clearer. As now addressed in response to comments made by yourself and several other reviewers, we do not discuss aerial (or seet) calls at all in this paper and all comparisons are between whining-type calls (made towards brood parasites) and terrestrial alarm calls. Fairy-wrens, like many other birds, have aerial alarm calls, but they are not explored at all in this study as they are not relevant to the questions we are asking. We have now changed the definition of alarm call in the paper to explicitly refer to terrestrial alarm vocalization and have also added that our use of “alarm call” follows established convention (e.g. Rowley et al. 1965 Emu 64: 251-297) in these study species.

2. Following on the point above about mobbing calls, given that the context to produce the whining calls is learned, I was surprised not to see any reference to Curio's famous work on this topic for alarm calls generally. In an elegant series of lab experiments Curio showed convincingly that some young songbirds learn to associate alarm calls with a particular stimulus they are seeing when they hear other birds alarming. Songbirds were taught that honeyeaters are dangerous, and that dish detergent bottles are also dangerous, and if I recall correctly these specific patterns of enemy recognition were transmitted across future generations (young birds imprinting on harmless items then passed this on to their own offspring). This work is relevant to the subject of this study and may also be relevant to the issue of how unique the brood parasitic alarm calls are (or whether they are just a slightly modified version of more general mobbing calls. In some cases, might they be more generally mobbing calls that become associated with parasitism through learning? In this context is also interesting that some cuckoos mimic birds of prey (Welbergen and Davies) and I expect are mobbed as such.

RESPONSE: Thank you for pointing this out. Curio’s work was cited in an earlier draft but was (unfortunately) removed. It has now been re-added to the paper, and we note that a field-adaptation of Curio’s design was conducted in Feeney & Langmore (2013 Biol. Lett. 9: 20130443), which specifically examined (and demonstrated) social learning of brood parasitic cuckoos including an appropriate response towards them in individual superb fairy-wrens (i.e. whining vocalizations and aggressive physical mobbing).

Minor points by line numbers

43. Is this really analogous to language? If so, are all alarm calls and mobbing calls similar?

RESPONSE: We note that senior author Blasi is an evolutionary linguist whose work has shown that human languages contain innate predispositions between pairing particular sounds with particular meanings (Blasi et al. 2016 PNAS 113: 10818-10823), which overturned long-standing convention in the field of linguistics. Our study extends on this finding by showing that this phenomenon also exists across hosts of brood parasites in their vocalizations towards brood parasites (a phenomena that is intuitive, but we cannot find instances where it has been quantitatively demonstrated), and in addition to showing it exists, we provide evidence that it provides a fitness benefit. That this vocalization also contains acoustic features that have been shown to elicit an innate danger response (e.g. Aubin 1991 Behav. Process. 23: 103-111), despite its production in the correct context being learned (e.g. Feeney et al. 2013 Biol. Lett. 9: 20130442), makes it the first vocalization that we know of that has been demonstrated to contain both learned and innate components. This makes it a unique intermediary between learned and innate vocalizations, which has implications for evolutionary linguistics as it helps reconcile the relationship between these vocalization types (e.g. see Számadó. & Szathmáry 2006 Trends Ecol. Evol. 21: 555-561; Wheeler & Fischer 2012 Evol. Anthro. 21: 195-205). Also, regarding the similarity between call types (whining-type and terrestrial alarm calls), we show that they are structurally distinct (see Evolutionary history of a sound-meaning association, and Fig. 1B).

53. Jargon. What does innate vocal reaction mean? What are some examples.

RESPONSE: This is a good point, thank you. We have modified the first paragraph to include an example (scream in response to pain) to better explain what we mean.

56. If it is arbitrary then why would we see convergence. I do not think this makes sense.

RESPONSE: There has been a long-standing belief that human languages are separated from animal communication systems due to them being characterized by learned-arbitrary symbols, with exceptions (e.g. onomatopoeia) being largely disregarded (see Wheeler & Fischer 2012 Evol. Anthro. 21: 195-205). However, it is being found (e.g. Blasi et al. 2016 PNAS 113: 10818-10823) that these exceptions are more widespread than they “should” be. This opens up the opportunity to explore why these phenomena might exist (i.e. we show they exist and confer a fitness benefit) and also offer an opportunity to try and better reconcile the evolutionary relationship between learned and innate vocalizations (which we do by describing a vocalization that contains both learned and innate components). We have changed the phrasing in this paragraph to help make our argument clearer.

I do not know about linguistics and ideas about human language to be able to make sense of these claims.

RESPONSE: We note that we have provided references in support of the statements we make throughout the introduction and discussion.

69. There is a mix of things here that is then summarized as a suggestion that does not make sense to me. Some aspects of words are shared (r in round) in human speech AND human screams and distress vocalizations elicit an innate recognition. Putting these two completely different things together " suggest that some aspects of behavioral responses in an otherwise learned system harness phylogenetically generalizable innate associations between sounds and meanings, perhaps due to them delivering an adaptive benefit". I do not understand how these things combine to lead to the suggestion.

RESPONSE: This section does draw on the linguistics literature to provide examples of how human languages contain innate predispositions like those seen in animal communication systems. We have made several changes throughout this paragraph to make our statements clearer to the broad readership of this journal.

101. (already summarizes above but included here for clarity) I am very surprised that nothing is mentioned about mobbing because in some ways these calls seem structurally similar to mobbing calls. It is also interesting that some cuckoos mimic raptors and are likely to mimic a mobbing response. If these whining calls are in some ways related to mobbing calls or derived from them, then the story is somewhat different than presented.

RESPONSE: Yes, we intentionally veered away from that in this paper, as mobbing calls (alarm calls in our study) often do not actually elicit physical mobbing, rather a vigilance response (see analyses added to more explicitly demonstrate this following reviewer 2's suggestion). By contrast, whining-type calls do elicit extremely aggressive physical mobbing. Regarding their supposed structural similarity, we note that our analysis specifically examines this (see Evolutionary history of a sound-meaning association, and Fig. 1B), and we find that alarm and whining-type calls are structurally distinct. The wide variation in alarm calls (or mobbing calls) that we observe (see Fig. 1A) is also evident in published literature (e.g. Ficken & Popp 1996 *Auk* 113: 370-380) where they conclude that the variation in "mobbing" calls is not likely a result of convergent evolution. To address this comment, we have added a brief discussion of our findings regarding the variability of alarm vocalizations, and have noted that it reflects the findings of past literature (e.g. Ficken & Popp 1996 *Auk* 113: 370-380).

If whining calls are basically mobbing calls, or are even modified mobbing calls, then the notion that whining calls are "functionally referential and harness an innate bias between sound and meaning" does not seem totally correct to me. That statement implies that all of these patterns and effects are exclusive to the specific context of brood parasitism. The effects of enemy recognition more broadly (predators) which is widespread among birds, needs to be parsed out.

RESPONSE: As explained in detail in response to Reviewer 2's first comment, "mobbing" calls (termed "alarm" calls in this manuscript) do not actually elicit mobbing, rather a more cautious vigilance response (see newly added statistics, following Reviewer 2's suggestion). They are also structurally distinct from whining vocalizations (Feeney et al. 2013 Science 342: 1506-1508), which are made in response to the sight of a brood parasite and do elicit extremely aggressive physical mobbing. Please see our response and corresponding changes already made in response to Reviewer 2's first comment.

135. "...contain conserved acoustic features that are known to elicit an innate danger response in other species". I am not sure what is being said here.

RESPONSE: We have now clarified this statement. Vocalizations with high bandwidth and low spectral flatness, such as distress calls, have been shown to elicit strong innate responses across taxa (e.g. Aubin & Bremond 1992 Behaviour 120; Aubin 1991 Behav. Proc. 23: 103-111).

156. Were the calls that were given to cuckoos versus predators analyzed as was done earlier in the paper to confirm that these birds are producing calls distinct for brood parasitism and that these differ from mobbing calls given to predators?

RESPONSE: This was not possible, as studies do not generally provide recorded examples (other than occasional exemplar spectrograms) of the vocalizations they analyze, thus these additional studies explored to see if they noted a unique vocalization towards brood parasites vs predators (working on the assumption that if a different vocalization is being made towards one versus the other, the author would [should] have noted this - e.g. as done in Payne et al. 1985 Behaviour 94: 108-126; Langmore et al. 2012 Behav. Ecol. 23: 798-805; Noh et al. 2021 Anim. Behav. 174: 51-61). Unfortunately, we can only work with what is provided in previously published studies, and we are confident that our approach is thus both conservative and appropriate. That said, we have added more information to the methods to highlight the scope of this analysis more appropriately.

A couple of the sounds from Xeno Canto seem potentially problematic. For *Prinia rufescens* there were several recordings of birds giving alarm calls and many sounded very different from each other. How did the authors decide to choose the example they chose. And for the *Philemon*

corniculatus example, the bird seems to be giving a regular call (and there is no mention of alarm call in the recording description).

RESPONSE: We appreciate the reviewer's attention to this. For the *P. rufescens*, we had chosen a single high-quality sound file from a predator context recording from a location close to where the parasite context calls were recorded. However, we agree with the reviewer that the alarm calls of the *P. rufescens* are variable across different recordings. As such, we have now included 4 distinct high quality predator context call recordings. For the *P. corniculatus* recording, there was only a single recording available on xeno-canto at the time the analysis was conducted. Following the reviewer's comment, we revisited the Xeno-Canto sound library and found an additional recording of calls described as alarm calls. These calls are now included in the updated analysis and manuscript.

Response to reviewer comments “NATECOLEVOL-24102745”

Reviewer #2 (Remarks to the Author):

I think that the additional detail provided by the authors strengthens the argument that whining calls are functionally referential to brood parasites and learned, which was my primary concern in my initial review. I do still have some minor concerns listed below, but if the authors can address these then I recommend that the paper be published. I think it is a fascinating and important study and would like to see it get the attention it deserves in a high-impact journal like *Nature Ecology and Evolution*.

RESPONSE: We appreciate the thorough review provided by this reviewer previously, and thank them for their additional minor comments. We have addressed them all, and are confident that they have helped us produce a clearer and more considered manuscript.

1. I appreciate the addition of an analysis to assess if birds exhibit a qualitatively different response to whining calls vs. terrestrial alarm calls. However, the response variable chosen (number of individuals approaching 0-2 meters vs. 2-5 meters) seems a little contrived to me, as if you were trying to artificially convert a continuous variable into a categorical response. For many of the host species you tested, I understand that there may not have been any other clear response to measure. However, In the case of yellow warblers, why not record whether or not the warblers flew back to their nest in response to the playbacks? That seems like the most obvious response variable to measure, since it apparently represents a clear qualitative difference in how yellow warblers respond to brood parasites vs. predators (line 154). Was there a reason you did not include this response variable?

RESPONSE: We apologise if our previous response gives the impression that we were attempting to convert a continuous variable into a categorical response, and we confirm that this was not the case. To clarify, our data collection protocol included recording whether birds approached within different distance categories, as these categories provide different information about the perceived threat each treatment presents. Regarding your suggested analysis using the yellow warblers, we unfortunately did not collect these data; however, we are also not convinced that it would provide useful information pertaining to this specific point (see response to comment 4, below). Overall, we are confident that our analysis addresses the reviewer’s request to provide additional evidence to show that hearing whining vocalizations elicits a qualitatively (in addition to quantitatively) different response to hearing terrestrial alarm vocalizations.

2. The description of the results for the playback experiment with yellow warblers is a little unclear. The wording of lines 542-544 makes it sound like the lack of a differential response to the

treatments in yellow warblers contradicts your hypothesis that whining calls elicit a universal innate response in passerines. But if I understand correctly, the fact that yellow warblers responded equally to fairy wren whining calls and to yellow warbler seet calls actually supports the hypothesis that responses to whining calls are innate and universal. Can you please clarify in the text whether the responses of yellow warblers to the playbacks supports or contradicts your hypothesis?

RESPONSE: Thank you for pointing out that we could have been clearer here. As indicated in text, our analyses show no difference in conspecific responses across treatments, but do show a difference in heterospecific responses. While no difference between the different treatments is consistent with a universal response, the lack of a difference from them and the control makes us cautious to draw meaning. We do note that we did identify a difference in our analysis of heterospecific responses (including no heterospecifics being recorded during control playbacks), which is consistent with an innate response in species that exist in areas where whining vocalizations have never previously been reported. We have added additional text highlighting our reservations about drawing meaning from the yellow warbler response.

3. In Figure S2 it looks like yellow warblers responded equally strongly to whining, chip, and seet playbacks, but much less strongly to control playbacks, yet line 543 of the main manuscript seems to indicate that there were no differences in response to any of these treatments. Can you clarify if indeed there was no statistically significant difference in response between the control playbacks and the other playbacks?

RESPONSE: The text is correct, while more yellow warblers approached during the whining, chip and seet vocalizations (as illustrated in Fig. S2), our analysis detected no significant difference between them and the control. This comment highlights that we could have been clearer, so we have changed the second half of this sentence to use more consistent wording between the conspecific response versus the heterospecific response.

4. I think you should provide a little more clarification about what you mean when you say that the response to whining calls is innate. Yes, you found that whining calls elicit approach from passerines with no prior exposure to whining calls. But this doesn't necessarily mean that these species understand whining calls as specific references to brood parasites. Rather, whining calls may simply elicit a heightened but non-specific response in naïve species by taking advantage of a pre-existing auditory bias in passerines. To determine if naïve species understand whining calls as referential labels for brood parasites, you would need to do some additional experiments. For example, if yellow warblers fly back to their nest in response to both seet calls and whining calls but not in response to chip calls, then this would imply that they innately treat whining calls as referential signals for brood parasites (which would be truly astounding!). Similarly, if naïve species respond to whining call playbacks in the breeding season but not outside the breeding

season, that might imply that they innately understand the calls as references to a threat that is only relevant during the breeding season (although you would need a control to ensure that the birds are not simply less responsive to all warning signals outside of the breeding season). I don't expect you to do these additional experiments for the current study, but I think you should provide some clarification in the manuscript about exactly what you mean when you say that the calls elicit an innate response. If naïve species are innately sensitive to the acoustic properties of whining calls but don't understand them as referential signals, would these calls still represent a bridge between innate-reactionary and learned-symbolic signals?

RESPONSE: This is an interesting comment that touches on several issues.

First, while the brown headed cowbird (*Molothrus ater*) is indeed a brood parasite, it is ecologically very different from most brood parasite species, and thus direct comparisons between their host(s) and hosts of most other brood parasites may be problematic. This is primarily because brood parasitism by the brown-headed cowbird does not typically result in the destruction of the host's entire brood, and thus some of the host's own chicks can be raised alongside the cowbird. This contrasts with most other brood parasite species (including all towards which we reported evidence of whining vocalizations) where successful parasitism results in the complete destruction of the host's own brood. This is very important, as detection of the brood parasite by the host does not necessarily matter that much for the brown headed cowbird - if a brown headed cowbird successfully lays an egg and the host rejects it from the nest, the cowbird can simply destroy the host nest until it accepts the cowbird's egg and raises it alongside some of its own chicks (i.e. 'mafia behaviour', see Hoover & Robinson 2007 *PNAS* 104: 4479-4483). However, such retaliatory mafia behaviour does not work for brood parasites that destroy a host's entire reproductive output, because the host receives no benefit from raising the parasitic offspring. Correspondingly, these more virulent brood parasites tend to rely very heavily on deception (e.g. secretive behaviours or mimetic phenotypes) to facilitate brood parasitism, and thus detection of them and effective information transmission about local parasitism risk is extremely important for facilitating host defences (e.g. Davies & Welbergen 2009 *Science* 324: 1318-1320; Thorogood & Davies 2012 *Science* 337: 578-580). Given this, while we may expect selection for a highly detectable are directional referential vocalization in hosts of virulent brood parasites, as it facilitates the effective transmission of information that delivers a clear fitness benefit via an ecologically appropriate response, we may not expect such selection in hosts of non-virulent brood parasites (i.e. brown headed cowbird) as it is not ecologically appropriate.

Second, that whining vocalizations contain both innate and learned components is exactly what makes them a unique bridge between innate and learned symbols. This combination of features bears a resemblance to learned words in human language that exhibit innate

sound-meaning associations (see Blasi et al. *PNAS* 2016 113: 10818-10823) and is also consistent with Darwin's early speculations regarding the origins of spoken language.

To address this comment, we have provided additional text to highlight the ecological differences between brown-headed cowbirds and the brood parasites towards which we recorded evidence of whining vocalizations, in hopes that this clarifies that there is no reason to expect that hearing whining vocalizations should elicit anything other than the same innate recruiting response that we see in other species. We have also changed our last subheading title (before the discussion) to "innate recruiting response towards whining vocalizations...", and changed the final sentence of the second paragraph of that section to say "... whining vocalizations appear to elicit an innate exaggerated recruiting response in passerines globally."

Reviewer #2 (Remarks on code availability):

Unfortunately, the link to the Dryad data/code still does not work for me. When I clicked on the link I got "Error: DOI not found".

RESPONSE: We again apologise for this and will work with the Editor to ensure all data is available.

Reviewer #3 (Remarks to the Author):

The authors properly addressed all my comments and the MS is now much improved both in clarity and in statistical robustness. I have no further comments.

RESPONSE: We thank the reviewer for their insightful comments on the previous version of the manuscript and their positive appraisal of our responses to them.

Reviewer #4 (Remarks to the Author):

I thank the authors for this substantial and thoughtful revision. They have fully addressed my previous comments, and their responses to the other reviewers' concerns are likewise satisfactory.

The study provides good evidence for a previously undocumented link between innate and learned vocalizations, representing a kind of missing link in the evolutionary history of language. The tantalizing implications for the evolution of communication systems, including the foundations of language, are further strengthened by the cross-continental scope of the experimental work, which significantly enhances the generality and impact of the findings.

I believe Darwin would have been most pleased with this elegant demonstration of how innate and learned elements may co-vary in communication systems.

Sonia Kleindorfer

RESPONSE: We are very appreciative of Prof. Kleindorfer's helpful comments on the previous version of the manuscript and their positive appraisal of our responses to them. We are confident that our manuscript is better because of their help.

Reviewer #5 (Remarks to the Author):

The authors have done a very thorough job of responding to all of the reviewers' comments and it is clear that they have an extremely interesting study with a wealth of experimental and observational lines of evidence to support their claims. I also like the very clear organization, with subsequent sections each clearly addressing a key question or component.

That said, I am still not sure that the authors have completely addressed a key concern I raised, a concern also raised by other reviewers. The issue is whether this whining call is distinct from the general mobbing calls that many species use to recruit other birds to mob a predator. At least in northern hemisphere songbirds, there are two classic "alarm" calls—a high pitched pure tone "seet" call that is used in response to a flying raptor. This results in birds freezing or diving to cover and is not associated with mobbing. These calls are also hard to locate. All of this makes sense because the predator poses an immediate threat to the lives of the birds giving the alarm call. In contrast, classic mobbing calls are broad frequency, often harmonic calls that are used to recruit other species to mob a predator. The structure of these calls makes them easy to locate—this helps in other mobbers join in the mobbing of the predator. These calls are used when there is a perched predators (or snake on the ground), which are much less threatening to the mobbers' survival than is flying predator. In some cases birds approach the model but are not aggressive towards it; in other cases, mobbers will actually attack the predator (hit the raptor on the head from behind). This is relevant the authors' point that hosts will not physically attack some birds (currawongs) but will attack others (cuckoos). This distinction (attack versus just come close) may not be useful to distinguish among types of alarm calls.

With this in mind, I am still a bit confused about the terminology the authors use and I am not convinced they have fully dealt with this issue in their response to reviewers. They refer to "terrestrial alarms". Are these used to recruit mobbers or do they function like the high frequency seet calls I mentioned above: the birds freeze or hide? Put another way, for the authors' claims to be correct, the birds with whining calls would also need to have a second call that also recruits mobbers, but for something other than a brood parasite. Alternatively, if the second alarm call functions like the seet alarm call then the claims made by the authors are overstated. In that case,

the whine call of the hosts could simply be a classic mobbing call that is learned to be used towards cuckoo, and their only real mobbing call. (As Curio's experiments showed, birds could learn that this mobbing call is for cuckoos and not other birds like owls or sitting raptors). In this case, we would not be talking about the evolution of a new alarm call type (whining) but simply using the mobbing type calls that many songbirds have for a specific enemy. And if these are just typical mobbing calls, and if these hosts do not have other mobbing calls but more high threat alarm calls, then it would not be remarkable that they differ in structure from the other alarm calls. An analysis of mobbing versus seet calls would show exactly the same thing. I raised this point in my initial review but despite the voluminous response from the authors to me and the other reviewers who raised similar issues, I am still not sure that I have a clear understanding of what the terms alarm calls and terrestrial alarm calls. Perhaps I am missing something. If the authors have clearly identified a call (whine) that is different and distinct from the more classic mobbing call (i.e. call used to recruit other birds to mob a sitting predator of some sort (typically bird but also snakes in some birds) while at the same time these species have a second different mobbing call for sitting predators that is used to recruit mobbers (i.e. not a call to indicate an immediate threat that is not used to recruit mobbers, then I think the paper is very novel and should be published. However, if these species have only one call that recruits mobbers, and this is used for cuckoos, and they lack a second different call to recruit mobbers for other threats, then the paper is not exactly what it claims to be. I ask that the authors provide a brief clear response to this question so that I can fully understand what they are showing here and what is unique and novel.

One thing that may be confusing me, and this study in general, is that northern and southern hemisphere birds may have different sound structures and uses of alarm calls. If this were the case, using the same labels for all of these species. Note that the prinia and the Old World warbler both have what looks like a pure tone "alarm" vocalization that may function as the standard "flying predator warning" rather than a recruitment for mobbers.

RESPONSE: We appreciate the reviewer's positive appraisal of our paper and our responses to their helpful comments.

We believe the continued confusion here stems from the fact that we do not discuss vocalizations towards aerial predators (termed "seet" vocalizations by the reviewer, which is in and of itself a source of confusion given that "seet" is also the name used to refer to the referential vocalization that North American yellow warblers make exclusively towards brown-headed cowbirds [Gill & Sealy 2004 *Behav. Ecol. Sociobiol.* 56: 71-80]) in this manuscript, as these vocalizations are not relevant to this study. However, we acknowledge that aerial alarm vocalizations exist, and highlight that Magrath et al. (2008) *Proc. R. Soc. B* 276: 769-774 examines aerial alarm vocalizations in several species, including both of our primary study species in Australia (superb fairy-wren and white-browed scrubwren), which may be of interest to the reviewer.

Hopefully this clarifies to the reviewer that (among other things) our study shows that 21 species from all around the world have a unique vocalization that they make towards brood parasites (whining vocalization) and that this vocalization is structurally different to that produced towards terrestrial predators (what the reviewer refers to as ‘mobbing’ vocalizations, see Fig. 1A and S1) and aerial predators (not analyzed in this study as they are not relevant to this study). We also note that our experiments show that hearing a terrestrial alarm/mobbing vocalization does elicit recruiting behaviour (e.g. Fig. 2G-J); however, it is a much more cautious recruiting response than that elicited by hearing whining vocalizations (Fig. 2A-B).

Regarding the old world warbler and prinia vocalizations the reviewer refers to, they are also terrestrial alarm/mobbing vocalizations and we highlight that our analysis of their structural variability reflects that previously observed in the literature (e.g. Ficken & Popp 1996 *Auk* 113: 370-380).

To avoid any further confusion, we have included a statement in the Vocalization dataset section of the methods stating that we do not include a discussion of alarm vocalizations towards aerial predators in this study.

Reviewer #6 (Remarks to the Author):

The authors use a combination of experimental and modelling approaches to demonstrate how learned vocalisations may derive from innate responses with shared evolutionary origin. The overall idea is fascinating and may be of interest to evolutionary biologists and linguistics. However, the fairly small sample sizes and inaccurate descriptions of statistical methods and results throughout the manuscript (see comments below) raise concerns about the scientific rigour of the study. The authors will need to improve their analyses and be much more transparent with their methods and results before the manuscript could be considered for publication.

Major concern:

One main conclusion of this ms is that whether a species produces whining calls (presence or absence) is associated with opportunities of inter-species cooperation (measured as linkage density). The authors have tried to argue that the two measures of cause and consequence are independent since data were collected from different sources (call type mainly from recordings and literature, and linkage density solely from literature), but the positive association may also be caused by the biases in previous knowledge. Common/better-studied species tend to have more complete information on parasite-host interactions and thus higher linkage density, have more xeno-canto recordings which make whining calls more detectable, and also are more likely to be documented as producing whining calls in literature – the last can be very problematic since nearly

half of the species were assigned their whining status (yes or no) based on literature alone. If the authors can repeat their models by limiting the samples to well-known species in well-studied areas (e.g. in North America/ Europe/ Australia) to reduce sampling bias and find similar results, the conclusion may be more convincing.

RESPONSE: We appreciate the reviewer's suggestion regarding potential biases in the positive association we observed between species that produce whining vocalisations and linkage density. Specifically, the reviewer highlights that better-studied species are more likely to be those that have been detected to produce whining vocalizations and exist in areas where linkage densities may be higher, due to them receiving more research attention. We thank the reviewer for raising this point and agree that addressing it in the text would be beneficial.

We fully agree that better-studied species are more likely to have detailed information available, and we note that we have previously assessed the geographic bias in brood parasitism research (Kennerley et al. 2022 *Ecol. Lett.* 25: 1889–1904) and found that it mirrors biases in the broader ecology literature (e.g. Amano et al. 2021 *PLoS Biology* 19: e3001296; Nuñez et al. 2021 *Trends Ecol. Evol.* 36: 766-769). What is fascinating, however, is that despite the skewed research bias towards North America and Europe (i.e. 70% of past field-based brood parasitism research has been conducted in these areas [$N = 892$ of 1353 studies, Kennerley et al. 2022 *Ecol. Lett.* 25: 1889-1904]), these areas exhibit low linkage densities, in a global context, and almost no evidence (i.e. none in North America and one in Europe) of species reported to produce whining vocalizations. By contrast, areas where comparatively less research has been conducted (i.e. Africa, Asia and to a lesser extent Australia) are those where linkage densities are highest (despite 15 brood parasites having no known hosts in Africa and Asia, Kennerley et al. 2022 *Ecol. Lett.* 25: 1889-1904) and where all other species that have been reported to produce whining vocalizations exist (23 species), according to data obtained from both recordings (21 species) and the literature (two species). This suggests if more comprehensive data on brood parasite–host relationships were available, such that regional biases were reduced, our reported relationship should become even stronger. As our result is contrary to that which would be expected if the results simply followed the biases of past research efforts, we do not see how running additional analyses would make our results more convincing. This comment highlights, however, that information regarding this point would be useful for future readers and we have added additional text, accordingly.

Line 184-185: '> 5 times more slowly in whining than in alarm vocalizations'? Is it alarm vocalisations or all vocalisation types put together? The LRT result suggested the latter. And what is 'LRT comparing models with and without vocalization-specific evolutionary rates' anyway?

There is no text about LRT models in methods so it's hard to tell. Please add relevant model details. And what does PC3 mean in line 185 of the LRT result?

RESPONSE: Thank you for highlighting that we could have been clearer here. The Adams (2013) function compares evolutionary rates of multiple traits (here, PC3 scores of whining and alarm vocalizations) on the same phylogeny. First, the method compares the likelihoods of a model with a common evolutionary rate for both traits to that of a model with separate evolutionary rates for each trait. Second, the model estimates the evolutionary rate(s) for both models. So, the LRT refers to the test performed by the Adams function, which compares models with and without distinct rates. A significant LRT implies support for distinct rates, which is what we found. In the model with distinct rates for each vocalization type, we compared the estimated rates, which were > 5 times slower for whining than for alarm vocalizations. In the previous version of the manuscript, we had written PC3 in parenthesis to indicate that the LRT referred to models comparing evolutionary rates of PC3 scores. In the updated version, PC3 is explicitly described in the first paragraph of the results.

Line 427-428: 'phylogeny explained much of the variation in species that produce whining vocalizations' – what is 'variation in species'? If you meant whether or not (a binary state) a species produces whining calls then please write it that way. Also, Pagle's lambda is not equal to how much variation explained by phylogeny (although this is common misconception), but only to what extent related species share similar traits to each other. So it's better to say 'producing whining vocalisations is phylogenetically conserved' or something similar.

RESPONSE: Thank you for this suggestion on how to present our results more clearly. We have amended the text following your suggestion.

Line 432: I don't think Fig. 1C shows anything about within-species cooperation.

RESPONSE: Thank you for highlighting this. We have removed reference to Fig. 1C.

Line 434-437: chaotic presentation of results. What do Bayesian model results say about the geographical distribution of whining species? I think the models only test whether or not a species produces whining calls. In line 437, how could Fig. 1C support model results of how many trees were supported? I understand the figure may corroborate the results in a way but citing a figure directly after phylogenetic model results suggests they convey exactly the same information.

RESPONSE: Thank you for pointing out that we could have been clearer in our presentation of these results. The results of the Bayesian models do not directly explain the geographical distribution of the whining call. Given the relationship between species that

produce the whining call and linkage density, we inferred a relationship with geographical distribution due to global patterns of brood parasite—host linkage density. We have now amended the text to improve clarity including mention of geographical distribution of the whining call and added an additional supplementary figure (Fig. S2) which presents the distribution of posterior means when modelling the relationship between the production of whining calls by a given species and its median brood parasite—host network complexity across its distribution measured by linkage density (the mean number of links per species within a given network).

Line 438-442: In my view Fig. 1c is probably more suitable for supporting these results, only.

RESPONSE: We agree and have changed the text to “(see Fig. 1C)” to help indicate that this is only meant to support the results.

In methods, when you refer to R functions please use quotation marks around the function names (or italicise) and give citation or a link to every package. For example, line 655 ‘we conducted a principal components analysis using princomp on all species and vocalizations’ is poorly written. I know that princomp is a basic R function for PCA but writing it as if it was a regular word is inappropriate. Please change all relevant cases throughout the ms.

RESPONSE: Thank you for pointing this out. We have clarified the function names using italics and specified the package.

Line 683: ‘observed and expect counts’ of what vocalisation type, whining or alarm? I assumed the former given the context but then it contradicts ‘greater than 2 was considered... less often than expected as whining vocalisations’. Is it actually less often or more often?

RESPONSE: We have now provided additional clarification that this analysis was performed on brood-parasite associated calls. We also appreciate the reviewer’s request for clarification about the interpretation of residuals, which was worded incorrectly in the previous version. For each species, residuals of classifications of brood-parasite associated calls were calculated for both whining and alarm call categories. Residuals less than -2 for the whining call category imply that fewer brood-parasite associated calls than expected get classified correctly.

Line 689: linear mixed effects models, not ‘linear effects models’.

RESPONSE: Thank you for pointing out this typo. We have fixed it.

Line 694: vague method, unclear how the whining status was defined using emmeans. E.g., ‘differed in post-hoc tests’, differed in terms of different vocalization contexts (brood parasite-associated vs alarm)? I couldn’t tell. If the linear discriminant scores ‘capture the acoustic features of vocalizations along the axis separating parasite-associated and alarm vocalizations’, the scores must be higher or lower (i.e., uni-directional) for a species to be considered whining, not just being different. From your responses to other reviewers, it sounded like you also included an interaction term between species and calling context, this is crucial for understanding the method but it is not mentioned anywhere in methods. Please specify the methods.

RESPONSE: We have now clarified that the model included an interaction term between species and vocalization context. We have also clarified that a significant difference between vocalizations used in the different contexts is, alone, insufficient to conclude that a species has whining vocalizations. We considered a species to produce distinctive ‘whining’ vocalizations when two conditions were satisfied in post-hoc tests: 1) their brood parasite-associated vocalizations were estimated to have negative linear discriminant scores, similar to those of the superb fairy-wren, Hume’s leaf warbler, and the greenish warbler, and 2) the difference between their linear discriminant scores of their brood parasite-associated and alarm vocalizations was significantly different from 0 and negative.

Line 768: What was the prior used? What was the thinning rate and the final sample size of models? These are important for Bayesian models and I couldn’t find any information. The model results are also presented very vaguely. They are only mentioned once in-line without much detail. A table detailing the sample size, effect size and CI of predictors, etc. is needed.

RESPONSE: We used the default flat priors set by the threshBayes function and have now added wording in the manuscript to address this along with the thinning rate and final sample size. We have also added an additional data file (Data S6) which details effect size, credible intervals and ESS for each model across 1,000 randomly sampled trees based on the Ericson backbone and 1,000 randomly sampled trees based on the Hackett backbone.

Line 753: ‘As host status for a given species may change through time, we only use documented host species in this analysis’ – I don’t understand what it means.

RESPONSE: We have removed this sentence for clarity.

Other comments:

Abstract: ‘..., indicating that convergent use facilitates cooperative defenses’. Adding ‘across species’ at the end may improve clarity given it is said just a few lines above ‘species which produce this vocalization are rarely socially cooperative’.

RESPONSE: Added.

Figure 1A. The caption only mentioned species with and without whining calls identified ($n \sim 50$), which are far less than the number of tips and families shown in the phylogenetic tree. I assume the tree shows all passerine families and the tips are all passerine species ($n \sim 5k$)? Please explain what the tree consists of, and add sample sizes for how many species are identified to produce whining/seet calls, how many are not. There also seems to be a purple/grey bar above the warbler which is different from whining/seet/no unique vocalisation, what call type is it? The caption of fig. 1C is also very uninformative about how the linkage density was calculated. Just citing one reference is not enough, please add details.

RESPONSE: Thank you for pointing out that we could have been clearer here. Yes, the underlying tree comprises Passerines (5085 species). We included fifty two species as they met our inclusion criteria (see methods), comprising 23 that were identified to produce whining vocalizations and 29 for which there is no evidence of them producing whining vocalizations, including the yellow warbler which is the only known species to produce the 'seet' vocalization. These data have now been included in the caption. Regarding the purple bar, that is a typo included from a previous draft that we overlooked. It has now been removed and we thank the reviewer for pointing it out. We also agree that more information would be useful here, and have added more information about how linkage density was calculated in the caption. In addition to citing Kennerley et al. 2022 (*Ecol. Lett.* 25: 1889-1904) we have provided a reference to the methods, where we outline all methods in detail.

Figure 2. what species are involved in the panel K? Not explained in the main text or caption. Perhaps it's not key but without such information it's impossible to read a phylogram. A version with all species names added needs to appear in supplementary information at least.

RESPONSE: This is an excellent point, thank you. We have now added a supplementary figure (S3) that includes all species names, and have made reference to it in the caption of Figure 2.

Line 690: adding sample size for glmmBMT models will improve transparency and credibility of results (e.g., what was the total number of vocalisations/scores? Should be 25 species $\times$ ~ 40 per species?)

RESPONSE: We have now added this information in the methods. The dataset included 816 calls from 25 species, with between 9-40 calls per species and context.

Line 756-757: Kennerley et al.25 should be Kennerley et al.22? Check references.

RESPONSE: Thank you for pointing this out. We are not sure how this happened, as references were inserted using a reference manager (Zotero). We have now fixed these two references.

Reviewer #6 (Remarks on code availability):

The link to data and code is still not working. DOI not found.

RESPONSE: We apologize for this and will work with the Editor to ensure all data is available.

Response to reviewer comments "NATECOLEVOL-24102745B"

Reviewer #2:

The authors have addressed all my concerns, and I now recommend that this article be published. I only have one very minor proof-reading comment:

RESPONSE: Thank you for your comments on this version, and previous versions, of our manuscript. Your reviews have been very helpful and we believe our manuscript is better because of them.

In line 770, when you say "we THEN produced a hexgrid", it makes it sound like producing the hexgrid is a separate step from the previous sentence that talked about calculating the opportunity for between-species competition. But in fact, as far as I can tell, producing the hexgrid was part of the process of calculating the opportunity for between-species competition. So I suggest removing the word "then" in line 770.

RESPONSE: Changed.

Remarks on code availability:

The link to the code on Dryad still does not work. I have never been able to access the code throughout this review process.

RESPONSE: We are aware of the issue with data accessibility and note that we provided a Dropbox link that provided all data and code to the Editor with our last submission. We will continue to work with the Editor to ensure all data and code are available.

Reviewer #5:

Remarks to the Author:

I thank the authors for carefully addressing my remaining questions. I still have slight lingering doubts about whether the whining call is unique as opposed to being a more general mobbing recruitment call that has been learned to apply specifically to cuckoos. Perhaps that is a non-issue or is consistent with what the authors are claiming. Given the number of reviewers on this paper, if this were a big concern others would have raised it as well. Given this, I am happy to sign off on this for publication. It is a really interesting paper and it will stimulate further work on the topic.

RESPONSE: We appreciate the comments provided by this reviewer on all versions of the manuscript and are happy to hear that they are satisfied that all their comments have been adequately addressed.

Reviewer #6:

The authors have addressed all of my comments and have much improved the clarity of their

result presentation, methods, and manuscript text. I appreciate their effort of revising the manuscript and recommend its publication.

RESPONSE: We appreciate the comments provided by this reviewer and agree that our results and methods are clearer because of them.

Minor edit: in response to my previous comments, lines 445-447 were added to justify why knowledge bias may not affect result robustness. The author provided good logical argument in their response but the rationale was not fully explained in the ms, so it may be clearer to add some points to show readers the contrast between research effort and documented linkage density.

RESPONSE: We agree and have added more detail accordingly.