Peer Community In Ecology

Studying a rare behavior in a polygamous bird: male parental care in great-tailed grackles

Marie-Jeanne Holveck based on reviews by André C Ferreira and Matthieu Paquet

A recommendation of:

Folsom MA, MacPherson M, Lukas D, McCune KB, Bergeron L, Bond A, Blackwell A, Rowney C, Logan CJ. **Investigating the rare behavior of male parental care in great-tailed grackles** *(2020) In principle acceptance by Peer Community in Ecology of the version on 15 June 2020.* https://github.com/corinalogan/grackles/blob/master/Files/Preregistrations/gmalecare.Rmd

Open Access

Submitted: 05 December 2019, Recommended: 12 June 2020

Cite this recommendation as:
Marie-Jeanne Holveck (2020) Studying a rare behavior in a polygamous bird: male parental care in great-tailed grackles. *Peer Community in Ecology, 100054.* 10.24072/pci.ecology.100054

Published: 15 June 2020

The Great-tailed grackle (*Quiscalus mexicanus*) is a polygamous bird species that is originating from Central America and rapidly expanding its geographic range toward the North, and in which females were long thought to be the sole nest builders and caretakers of the young. In their pre-registration [1], Folsom and collaborators report repeated occurrences of male parental care and develop hypotheses aiming at better understanding the occurrence and the fitness consequences of this very rarely observed male behavior. They propose to assess if male parental care correlates with the circulating levels of several relevant hormones, increases offspring survival, is a local adaptation, and is a mating strategy, in surveying three populations located in Arizona (middle of the geographic range expansion), California (northern edge of the geographic range), and in Central America (core of the range). This study is part of a 5-year bigger project.

Both reviewers and I strongly value Folsom and collaborators' commitment to program a study, in natural field conditions, of a rare, yet likely evolutionary-important behavior, namely parental care by males of the great-tailed grackle. Yet, we all also recognized that it is a risky endeavor, and as a consequence, we wondered about the authors' ability to reach a sufficient sample size to statistically test (all) hypotheses and predictions with enough confidence (e.g. risk of type I errors, also known as false positives).

Folsom and collaborators acknowledged these limitations in their pre-registration. (i) They made the exploratory nature of their research work clear to readers. (ii) They adapted their analysis plan in running prior power analyses and in focusing on effect sizes (estimates and confidence intervals). (iii) Last and not least, Folsom and collaborators clearly exposed a priori hypotheses, their associated predictions and

alternatives, and ranked the latter based on their plausibility according to knowledge in the current and other study systems. Developing theory about male parental care behavior more generally with regard to a polygamous species that is rapidly expanding its geographic range and that is considered not to provide male parental care is without any doubt an added value to this study.
In summary, while this study will likely be insufficient to fully understand male parental care behavior of great-tailed grackles, it is much needed because it will definitely allow rejecting some predictions (e.g., if this behavior is present in all the studied populations, it would be common across range against expectation; finding only one male providing care to an unrelated offspring would lead to reject the prediction that males only care for their own offspring) and thus it will help laying the foundation of future research directions.
I strongly support the pre-registration system and thank all the contributors for producing a fruitful discussion throughout the review process, which in fine improved the clarity and logic of this pre-registration. Given the positive and encouraging reviews, the detailed and thorough answers to all comments by Folsom and collaborators, and their satisfactory and interesting revisions, I am happy to recommend this pre-registration and I look forward to seeing its outcomes.

References

[1] Folsom MA, MacPherson M, Lukas D, McCune KB, Bergeron L, Bond A, Blackwell A, Rowney C, Logan CJ. 2020. Investigating the rare behavior of male parental care in great-tailed grackles. corinalogan.com/Preregistrations/gmalecare.html In principle acceptance by PCI Ecology of the version on 15 June 2020 corinalogan/grackles/blob/master/Files/Preregistrations/gmalecare.Rmd.

Revision round #2

2020-05-26
Dear Folsom MA, MacPherson M, Lukas D, McCune KB, Bergeron L, Bond A, Blackwell A, Rowney C, Logan CJ,

I sent your preprint entitled "Repeated parental care by adult male great-tailed grackles and its association with hormones, fitness, specific populations, and mating strategies" out for a second round of reviews. Indeed, while you gave satisfying and detailed answers to most comments, some points remained to be clarified. In particular, I was wondering if the study was really worthwhile to be carrying-out because, as you realized by yourselves, you probably won't reach sufficiently high sample sizes, with the proposed method for data collection, to statistically test your hypotheses and predictions.

The reviewers are the same as in the first round. Their comments are appended below. I largely agree with them. You should make the exploratory nature of your research work clearer to readers. Thus, it should clearly be mentioned in the abstract, together with the very low expected sample sizes. Now, the abstract sounds really promising (with a lot of hypothesis testing), while in the end likely no (clear) answers will come out. I also think that the suggestion of André Ferreira to include a section discussing the potential ways of improving the current method for data collection, and to present it as one of the outcomes of your project, is necessary to enforce the exploratory nature of your project. This should also appear in the abstract.

We therefore invite you to revise and resubmit your preprint, taking into account all of the points raised.

We look forward to hearing from you soon.

Best regards,

Marie-Jeanne Holveck, Recommender for PCI Ecology

Preprint DOI: http://corinalogan.com/Preregistrations/gmalecare.html

Reviewed by Matthieu Paquet, 2020-05-03 14:11

I thank the authors for taking the time to answer all my comments in a detailed way and I am happy with this version of the preregistration. I only have small additional suggestions and comments below if the authors wish to address them. I wish them all the best with their upcoming field seasons.

Regarding Response 4: I could not access the reviewer comments of the preregistered study mentioned as I guess it has not been recommended yet. I suppose that chicks cannot be ringed in the study species which is why males are not aged. Perhaps this could be mentioned here (or at least in the future manuscript that will result from this study)?

Regarding response 11: note that in Kerr 1998, the plausibility of a hypotheses is the second key factor to consider together with whether the hypothesis was anticipated or not to prevent "HARKing". I appreciate that the authors may not have clear ideas of the a priori likelihood of each hypothesis but I felt that based on previous literature (including monogamous species or non-avian species with facultative male care) we could expect different likelihoods regarding the existence and the sign of relationships between parental care, condition and territory holding. Note that my concern was not so much about stating too many hypotheses but to not state how plausible they seem based on knowledge in the current and other study systems.

Regarding power analyses: it would be helpful to clarify why the effect sizes in the input f^2=1.5 or 2.2 are different from the f^2 given in the interpretation of the results of the power analyses "This means that, with a projected sample size of 9, we have a 71% chance of detecting only very large effects (approximated at f2=0.80 by Cohen (1988))."

Regarding Response 16: "b) We realize that with our sample size we will only have the power to detect very large effects. Accordingly, we revised the preregistration to say: Analysis Plan > H1 P1-P3 and H4 mating strategy: "We will determine whether an explanatory variable had a detectable effect using the Estimate in the full model: if the 89% confidence intervals do not overlap zero, we will consider this a detectable effect (@statrethinkingbook).”" The idea behind the 89% credible intervals (if I interpret this excellent book correctly) is rather to give an interval that is large but different than the classical 95% (also chosen for an arbitrary reason: it is a prime number) in order to NOT be tempted to use it as a significance test. Given the small expected sample size (and hence low power), perhaps a better approach would be to forget about "detectable" or "significant" effects but simply focus on estimates and confidence intervals (95, 89 or whatever large value)?

Reviewed by André C Ferreira, 2020-05-25 18:46

I thank the authors for completing a thorough and complete revision based on the comments of the reviewers, especially on the clarification about how the behaviour observations are made, the sample size rationale and the exploratory nature of this project. I only have a few minor comments to add:

1)I understand that due to limitations of resources there is little that the authors can do at the moment to increase the opportunities to collect data on the interactions between adult male grackles and fledglings. However, since the authors are aware of this limitation and they are currently studying the possibility of using ICARUS satellite tags to automate this type of data collection, I think that this project should include a section discussing the potential ways of improving the current method for data collection. For example, the authors could mention that early trials of ICARUS satellite tags could potentially be run at the later stages of this project. Not only this development of new methods is clearly needed to collect enough data to study this rare behaviour, it would also help the readers to understand the exploratory nature of the project and would show the authors' commitment to improve the current method of data collection. This improvement in data collection could be presented as one of the outcomes of this project.

2)Since the sample size will be almost certainly small is it adequate to run power analysis, to mention the use of general linear models and that the data will be checked for overdispersion, underdispersion, zero-inflation, and heteroscedasticity?

3) Regarding response 19 point 4): “There are a few reasons we believe that our other data collection measures (above) are sufficient to detect male parental care based on what we have found in Arizona: 4) The males we have observed providing care do so repeatedly (apparently daily), therefore, while this appears to be a behavior that is rare among males, once a male expresses it, he does so regularly, therefore making his behavior more detectable.”. I must warn the authors that it is premature with the current data to make such assumption. The authors might have the impression that males that do this behaviour do it often, not because this is true but because the current method might be insufficient to observe males that exhibit this behaviour at a lower rate.

As in my first review I find this project very interesting and I value the authors’ commitment to study a rare behaviour. I still think that studying this behaviour in detail would require a different approach that would allow collecting more data than the one that is currently presented. Nevertheless the authors are open about this limitation and are open about the exploratory nature of this project and therefore this should not be a reason for rejection. I believe that this project will most certainly be of great interest to many researchers studying parental care behaviour.

Author's reply:

Dear Dr.’s Marie-Jeanne Holveck, Matthieu Paquet, and André Ferreira,

Thank you very much for your time in evaluating our revised preregistration and for offering us the opportunity to revise and resubmit. As before, we found your feedback helpful and useful and we are happy to conduct another revision. We revised our preregistration and associated files at http://corinalogan.com/Preregistrations/gmalecare.html, and we responded to your comments below.

Note that the version-tracked version of this preregistration is in rmarkdown at GitHub: https://github.com/corinalogan/grackles/blob/master/Files/Preregistrations/gmalecare.Rmd. In case you want to see the history of track changes for the preregistration at GitHub, click the previous link and then click the “History” button on the right near the top. From there, you can scroll through our comments on what was changed for each save event and, if you want to see exactly what was changed, click on the text that describes the change and it will show you the text that was replaced (in red) next to the new text (in green).

Thank you again for the opportunity to share our work with you.

All our best,

Melissa, Maggie, Dieter, Kelsey, Luisa, Angela, Aaron, Carol, and Corina

Investigating the rare behavior of male parental care in great-tailed grackles

Folsom MA, MacPherson M, Lukas D, McCune KB, Bergeron L, Bond A, Blackwell A, Rowney C, Logan CJ

http://corinalogan.com/Preregistrations/gmalecare.html version 3.0

Submitted by Corina Logan 2019-12-05 17:38

Abstract

This is a PREREGISTRATION submitted for pre-study peer review. Our planned data collection START DATE is May 2020, therefore it would be ideal if the peer review process could be completed before then.

Abstract: Great-tailed grackles (Quiscalus mexicanus) are known to have a mating system where females are the sole builders of the nest and caretakers of the young (Johnson et al. (2000)). However, there is one report from Selander (1970) in which an unmarked male great-tailed grackle provided parental care to two juveniles

in Austin, Texas. Here, we report repeated parental care events by multiple individually marked adult males in a population of great-tailed grackles in Tempe, Arizona. We also report male parental care by unmarked adult males in Tempe, Arizona; Santa Barbara, California; and San Clemente, California. We then propose to test hypotheses to determine whether male parental care 1) is associated with hormone profiles (testosterone, estrogen, prolactin, progesterone) potentially because hormones mediate investment in competition and care behavior; 2) increases direct fitness; 3) is a local adaptation, unique to specific populations where it has evolved; and 4) is a mating strategy. Keywords: birds, great-tailed grackle, parental care, male care, hormones, fitness, adaptation, mating strategy

Round #2

by Marie-Jeanne Holveck, 2020-05-26 17:08

Manuscript: http://corinalogan.com/Preregistrations/gmalecare.html version 3.0

Ask for a revision

Dear Folsom MA, MacPherson M, Lukas D, McCune KB, Bergeron L, Bond A, Blackwell A, Rowney C, Logan CJ,

COMMENT 1 I sent your preprint entitled "Repeated parental care by adult male great-tailed grackles and its association with hormones, fitness, specific populations, and mating strategies" out for a second round of reviews. Indeed, while you gave satisfying and detailed answers to most comments, some points remained to be clarified. In particular, I was wondering if the study was really worthwhile to be carrying-out because, as you realized by yourselves, you probably won't reach sufficiently high sample sizes, with the proposed method for data collection, to statistically test your hypotheses and predictions.

RESPONSE 1: We can see your point. One of the ways our article adds value is by developing theory about this behavior more generally with regard to a polygamous species that is rapidly expanding its geographic range who is considered not to provide male parental care. Even though our data will likely be too limited to fully understand this behavior, it is likely that we will be able to reject some of our hypotheses (e.g., if both territory holding and non-territory holding males are found to engage in male parental care then there will be no support for prediction 6; if this behavior is present in all populations, then we can reject prediction 5; if we find even one male who provides care and is not the father of the offspring, then we can reject prediction 4; see Response 7 for details). We revised the preregistration to reflect this added value by revising as follows:

Abstract: "We then develop hypotheses about why this behavior might occur in a polygamous species that is rapidly expanding its geographic range, positing that male parental care 1) is associated with hormone profiles (testosterone, estrogen, prolactin, progesterone) potentially because hormones mediate investment in competition and care behavior; 2) increases the number of offspring that survive to independence; 3) is a local adaptation, unique to specific populations where it has evolved; and 4) is a mating strategy. Finally, we propose to begin to collect data that will contribute to investigating these hypotheses. Given the rare nature of this behavior, our sample size will likely not be large enough to fully investigate these hypotheses, however it is our hope that this data can serve as a motivator for more targeted data collection to occur across this species' range."

COMMENT 2 The reviewers are the same as in the first round. Their comments are appended below. I largely agree with them. You should make the exploratory nature of your research work clearer to readers. Thus, it should clearly be mentioned in the abstract, together with the very low expected sample sizes. Now, the abstract sounds really promising (with a lot of hypothesis testing), while in the end likely no (clear) answers will come out. I also think that the suggestion of André Ferreira to include a section discussing the potential ways of improving the current method for data collection, and to present it as one of the outcomes of your project, is necessary to enforce the exploratory nature of your project. This should also appear in the abstract. We therefore invite you to revise and resubmit your preprint, taking into account all of the points raised.

RESPONSE 2: We made the likely small sample size more explicit and we clarified that we are developing theory on this topic in the abstract (see Response 1). Regarding additional data collection methods, please see our Response 9 below. It is difficult to split our larger long-term research program into smaller pieces and to decide where the various pieces belong. We decided to include pieces that are not certain to be feasible in our current funding period in separate future preregistrations. Thank you very much for your invitation to revise and resubmit!

We look forward to hearing from you soon. Best regards, Marie-Jeanne Holveck, Recommender for PCI Ecology

Reviews

Reviewed by Matthieu Paquet, 2020-05-03 14:11

COMMENT 3 I thank the authors for taking the time to answer all my comments in a detailed way and I am happy with this version of the preregistration. I only have small additional suggestions and comments below if the authors wish to address them. I wish them all the best with their upcoming field seasons.

RESPONSE 3: Thank you so much for your really useful feedback! We are so glad that you are happy with the revision. Thank you also for your further suggestions and comments. We address them below.

COMMENT 4 Regarding Response 4: I could not access the reviewer comments of the preregistered study mentioned as I guess it has not been recommended yet. I suppose that chicks cannot be ringed in the study species which is why males are not aged. Perhaps this could be mentioned here (or at least in the future manuscript that will result from this study)?

RESPONSE 4: Sorry about that! You are correct: the other preregistration has not yet passed pre-study peer review, therefore the review history is not online yet. You are also right in that chicks can't be ringed because the grackles nest in very tall trees and palms, which makes it unfeasible to reach the chicks (without lots of health and safety approvals that were not worth the extensive time investment for obtaining in Tempe). We added a note about this in the revision:

Planned sample: "We are unable to age them beyond juvenile (<1 year) or adult (>1 yr), and it has been thus far unfeasible to color mark chicks in the nests (which would allow us to track their age) because the nests are inaccessible."

COMMENT 5 Regarding response 11: note that in Kerr 1998, the plausibility of a hypotheses is the second key factor to consider together with whether the hypothesis was anticipated or not to prevent "HARKing". I appreciate that the authors may not have clear ideas of the a priori likelihood of each hypothesis but I felt that based on previous literature (including monogamous species or non-avian species with facultative male care) we could expect different likelihoods regarding the existence and the sign of relationships between parental care, condition and territory holding. Note that my concern was not so much about stating too many hypotheses but to not state how plausible they seem based on knowledge in the current and other study systems.

RESPONSE 5: Thank you for clarifying your original comment and for drawing our attention to the plausibility piece in Kerr 1998. When a prediction has alternatives, we expect the first prediction to be the most likely and the alternatives to be less likely. Where a prediction has no alternatives, we are unsure of which is more likely than the other. In the latter case, we can see where plausibility would be a useful piece to include, however given our experiences with the extensive population differences in this species (e.g., there are huge site differences in how easy they are to catch and with what methods, how willing they are to approach new objects and participate in experiments, how heavy they are, and how large they are) and with this species' tendency to provide unexpected results, we still think that we are unable to speculate about plausibility in this case. For an example of a recent unexpected result, I'll share a preview of something we just discovered: contrary to most other bird species (which are primarily monogamous), the males are the dispersing sex in this species (the post-study write up will appear soon at Sevchik et al. 2019

http://corinalogan.com/Preregistrations/gdispersal.html). To try to clarify this discussion, we made the following revisions:

Hypotheses > "Note that the predictions we consider more likely are listed as the main prediction and the less likely options are listed as alternatives. When we are less certain of likelihood, we list all predictions as main predictions."

Hypotheses > Prediction 3: we removed the Prediction 3 title and change it to "Prediction 1 or 2 alternative (independent)"

COMMENT 6 Regarding power analyses: it would be helpful to clarify why the effect sizes in the input f^2=1.5 or 2.2 are different from the f^2 given in the interpretation of the results of the power analyses "This means that, with a projected sample size of 9, we have a 71% chance of detecting only very large effects (approximated at f2=0.80 by Cohen (1988))."

RESPONSE 6: Good point. We clarified the text as follows:

Analysis plan > H1: "@cohen1988statistical approximates small effect sizes at f^2^=0.20, medium at f^2^=0.50, and large at f^2^=0.80. This means that, with a projected sample size of 9 and a target 70% chance of detecting an effect, we will only be able to detect very large effects (f^2^>2.2)."

Analysis plan > H1: "@cohen1988statistical approximates small effect sizes at f^2^=0.20, medium at f^2^=0.50, and large at f^2^=0.80. This means that, with a projected sample size of 9 and a target 70% chance of detecting an effect, we will only be able to detect very large effects (f^2^>1.5). "

COMMENT 7 Regarding Response 16: "b) We realize that with our sample size we will only have the power to detect very large effects. Accordingly, we revised the preregistration to say: Analysis Plan > H1 P1-P3 and H4 mating strategy: "We will determine whether an explanatory variable had a detectable effect using the Estimate in the full model: if the 89% confidence intervals do not overlap zero, we will consider this a detectable effect (@statrethinkingbook)."" The idea behind the 89% credible intervals (if I interpret this excellent book correctly) is rather to give an interval that is large but different than the classical 95% (also chosen for an arbitrary reason: it is a prime number) in order to NOT be tempted to use it as a significance test. Given the small expected sample size (and hence low power), perhaps a better approach would be to forget about "detectable" or "significant" effects but simply focus on estimates and confidence intervals (95, 89 or whatever large value)?

RESPONSE 7: Thank you for this suggestion - it makes sense given the nature of what will likely be a small sample size. Instead of making a subjective rule about what we are able to detect, we can instead discuss the sizes of the estimates, the location of their confidence intervals, and include what we are able to detect as part of this discussion. We removed "detectable" from this section and instead discuss how the estimates relate to the predictions. We revised as follows:

Analysis plan > H1: "We will reject prediction 1 if prediction 2 is supported. We will identify whether there is support based on whether estimate for T ~ caring is positive and the 89% confidence intervals do not overlap zero (@statrethinkingbook). If there is the estimate for T ~ caring is negative and the 89% confidence intervals do not overlap zero, then we will consider prediction 1 supported and prediction 2 rejected. If the 89% confidence intervals do overlap zero, we cannot be certain whether our sample size was sufficient to reject predictions 1 or 2 and whether prediction 1 or 2 alternative was supported."

Analysis plan > H4: "We will reject predictions 6, 7, and 8 if both roaming and territory holding males provide care. We will consider a prediction rejected if the 89% confidence intervals overlap zero (@statrethinkingbook). We will reject predictions 9 if the 89% confidence intervals for the estimate of care ~ territory holding do not overlap zero (@statrethinkingbook), which suggests that there is a strong enough effect and a large enough sample size to detect a difference. We will reject prediction 10 if the 89% confidence intervals for the estimate of care ~ territory holding and care ~ condition do not overlap zero (@statrethinkingbook)."

Reviewed by André C Ferreira, 2020-05-25 18:46

COMMENT 8 I thank the authors for completing a thorough and complete revision based on the comments of the reviewers, especially on the clarification about how the behaviour observations are made, the sample size rationale and the exploratory nature of this project. I only have a few minor comments to add:

RESPONSE 8: Thank you so much for your positive feedback! We are happy to address your additional comments (see below).

COMMENT 9 1)I understand that due to limitations of resources there is little that the authors can do at the moment to increase the opportunities to collect data on the interactions between adult male grackles and fledglings. However, since the authors are aware of this limitation and they are currently studying the possibility of using ICARUS satellite tags to automate this type of data collection, I think that this project should include a section discussing the potential ways of improving the current method for data collection. For example, the authors could mention that early trials of ICARUS satellite tags could potentially be run at the later stages of this project. Not only this development of new methods is clearly needed to collect enough data to study this rare behaviour, it would also help the readers to understand the exploratory nature of the project and would show the authors' commitment to improve the current method of data collection. This improvement in data collection could be presented as one of the outcomes of this project.

RESPONSE 9: We wanted to write this preregistration so as not to come up with posthoc hypotheses for any of the data that we collect. This is one step in the larger project and we want to make sure we clearly delineate this step before analyzing the data. Because it is unclear whether we will be able to implement the ICARUS tags during our current funding period (which covers the current preregistration), we would like to keep this preregistration focused on what we will be able to achieve across all of the field sites. We would need to develop a different set of predictions to cover the data collected by the ICARUS tags because the type of data collected would be about the proximity of adult males to juveniles. We could answer questions such as: What percentage of the day are males in close proximity to juveniles and how close to they get? This data would allow us to detect potential male parental care when we are not observing these individuals, which would indicate which males we should focus observation effort on. We have plans to keep the grackle research going beyond the current funding period and we will write new preregistrations to cover future hypotheses around male parental care after the work in the current preregistration is finished and we know more about how we want to pursue this line of inquiry. In the post-study write up of this preregistration, we can certainly include a "Future Directions" section where we discuss the various options available to expand this research.

COMMENT 10 2)Since the sample size will be almost certainly small is it adequate to run power analysis, to mention the use of general linear models and that the data will be checked for overdispersion, underdispersion, zero-inflation, and heteroscedasticity?

RESPONSE 10: This is a good point. We removed the Data Checking section because we will have to make do with whatever data we are able to get. We want to keep the power analyses in to give an indication of the kinds of power we would have with a small sample size, which will help readers interpret our results. We also want to keep the GLMs in because in H1 we will analyze the data for both sexes in the same model so the overall sample size will be much larger because it won't be restricted only to the males who provide care, and in H4 we will analyze all banded males so this sample size will also be larger than only the caring males. We will discuss the results cautiously so readers know how to interpret what we find about the caring males based on how many we end up having data on (see Response 7).

COMMENT 11 3) Regarding response 19 point 4): "There are a few reasons we believe that our other data collection measures (above) are sufficient to detect male parental care based on what we have found in Arizona: 4) The males we have observed providing care do so repeatedly (apparently daily), therefore, while this appears to be a behavior that is rare among males, once a male expresses it, he does so regularly, therefore making his behavior more detectable.". I must warn the authors that it is premature with the current data to make such assumption. The authors might have the impression that males that do this

behaviour do it often, not because this is true but because the current method might be insufficient to observe males that exhibit this behaviour at a lower rate.

RESPONSE 11: That's a fair point - it is too soon to make such generalizations.

COMMENT 12 As in my first review I find this project very interesting and I value the authors' commitment to study a rare behaviour. I still think that studying this behaviour in detail would require a different approach that would allow collecting more data than the one that is currently presented. Nevertheless the authors are open about this limitation and are open about the exploratory nature of this project and therefore this should not be a reason for rejection. I believe that this project will most certainly be of great interest to many researchers studying parental care behaviour.

RESPONSE 12: Thank you very much!

Revision round #1

2020-03-30
Dear Folsom MA, MacPherson M, Lukas D, McCune KB, Bergeron L, Bond A, Blackwell A, Rowney C, Logan CJ,

Your preprint entitled "Repeated parental care by adult male great-tailed grackles and its association with hormones, fitness, specific populations, and mating strategies" has now been reviewed and the reviewers' comments are appended below. You will see that, while they find your work of interest, they have raised points that need to be addressed by a revision. Both reviewers and I strongly value your effort to program a study, in natural field conditions, of a rare, yet likely evolutionary-important behaviour, namely parental care by males of the great-tailed grackle. Yet, we all also recognised that it is a risky endeavour, and as a consequence, both reviewers wonder about your ability to reach a sufficient sample size to statistically test (all) hypotheses and predictions with enough confidence (e.g. risk of type I errors, also known as false positives). For instances, Reviewer 1 thus suggests running prior power analyses to clarify this issue; Reviewer 2 suggests some automated methodologies to boost the sampling effort. Both reviewers also have a series of additional relevant comments, notably regarding the null hypotheses' identity, the rationale behind some of the predictions and behind the hypothesis 3, the age of the birds, the stopping rules for data collection, the statistical analyses, the conditions responsible for the occurrence of male parental care, the catchiness of the title, etc. (please see the detailed reviews). I believe that addressing their comments properly will highly improve your project.

We therefore invite you to revise and resubmit your preprint, taking into account all of the points raised.

We look forward to hearing from you soon.

Best regards,

Marie-Jeanne Holveck, Recommender for PCI Ecology

Preprint DOI: http://corinalogan.com/Preregistrations/gmalecare.html

Reviewed by Matthieu Paquet, 2020-02-04 11:14

The present pre-print aims to better understand the appearance and the fitness consequences of male post-hatching (post fledgling?) care, a very rarely observed behaviour in the study species, the great-tailed grackle (Quiscalus mexicanus). One important aim of the study is to look at whether such behaviour correlates with the circulating levels of several relevant hormones.
I thank the authors for taking the time to pre-register their study plan. I found the topic interesting and

stimulating. It is a particularly risky project since the study behaviour seems very rare (hence decent sample sizes may be hard to reach) but obviously worth investigating. I have some comments and suggestions that I hope may help the authors to get the best out of their project. (Feel free to contact me if some of the referred literature is not accessible to the authors)

A first general comment: It would be great to give some rough estimation of prior probability of the different predictions based on current knowledge of the study species and on what is known in other species. I am not saying that these should be all ranked according to a prior probability nor given any number. But it is important to state whether the different predictions seem likely or not, first to judge whether they seem worth testing or not and second to later on have an honest discussion on whether the results were expected or not and how likely they are to be false positive. (One can always write all possible predictions and then later on state that the results were predicted). Finding statistically "significant" results when testing very unlikely hypothesis is highly likely to be a false positive http://fpr-calc.ucl.ac.uk/.

A second general comment: Is there any information on the age of the birds? If so, it would be interesting to know if males providing care are essentially young or old individuals. This would give important information on the understanding of the system, notably depending on whether these individuals happen to be genetic parents, social parents, or alloparents (i.e. helpers). I am not saying that age should be included in the statistical models (there will certainly not be enough power to detect any age effect on the top of the rest and it is not a cofounding factor), but that it could be looked at in a more descriptive way.

Please find below some more specific comments/suggestions:

- Prediction 2 (additive): If high T (thus lower E, prolactin, and progesterone) Why lower prolactine in this case? Since prolactin is linked with parental care one would assume that males that provide care have higher T than males that do not have territories but also higher prolactine than males without territories and males with territories but that don't provide care?

- Prediction 3 (independent): Hormone levels do not systematically differ across individuals, regardless of sex or caring behavior, indicating that these hormones are not involved in parental care.
  Is this really predicted knowing prior knowledge on the role of these hormones in parental care in birds? It is important to test for a priori likely hypotheses in order to minimise the chance of false discoveries. This is rather the null hypothesis and I doubt it will be tested. I rather guess that the other hypotheses will be tested against this one.

- H2: Perhaps name it "relatedness" as local adaptation for hypothesis 3 sounds quite genetic as well.

- Prediction 4 (direct fitness): Males who provide parental care increase their direct fitness by caring for their own offspring. The more care offspring receive, the more likely they are to survive, therefore any male care will have a direct fitness benefit.
  This prediction makes sense but note that this is not affecting their direct fitness. This is affecting their offspring's fitness (and therefore affecting male inclusive fitness). Testing for direct fitness benefits would mean testing whether caring males survive better or increase their future number of offspring (egg laid), for example if providing care increases their attractiveness towards females, which is a possibility (e.g. Requena, & Machado 2015 and cf alternative 4). Of course it is always hard to draw a line between parents and offspring fitness in species providing post egg-laying care, but this is where the line is drawn in quantitative genetics anyway (e.g. Hadfield 2012, Smiseth 2012). To make is less problematic it would for sure be fine to just state that males increase their offspring fitness (instead of mentioning their direct fitness).

- Prediction 4, alternative 4 (signal): If the caring male is not the father of the offspring he is caring for, the caring behavior might be a signal to the offspring's mother, which then increases his chance of fathering her next offspring. Note that this may be true even if he IS the father. Females could still use this cue as a signal affecting whether or not they will breed with the same male again.

- Prediction5: Perhaps male parental care is more prevalent near the edge of their range because the costs of raising young might be higher (e.g., establishing in a new location, exposure to new predators, less protection due to a lower conspecific population density) and require a higher investment from parents.
  This sounds counter-intuitive. Higher cost of care should reduce the probability for parental care to occur (if benefits stay constant). What matters is the cost-benefit balance (See e.g. Lessells 2006, Alonso-Alvarez, C., & Velando, A. 2012).

- Prediction 10 (no relationship): Male parental care is not related to physical condition or whether they hold a territory.
  As for prediction 4, is that very likely that parental care is not related to physical condition or territory holding? These 2 factors surely affect the cost of care so they are predicted to affect the occurrence of parental care. This is rather a null hypothesis.
  When will blood sampling be conducted? Will they be sampled during both breeding and non breeding season? It would be great to have also blood samples of those food provisioning males when they are not breeding (and some non-caregiving males and females as well as control) in order to better understand the link between food provisioning and hormonal levels, also using intra-individual comparisons (do male care takers have high levels anyway or does their level vary depending whether they provide food or not?). Since we can expect a quite small number of male providing care, it would be a very precious additional information to get from them. In addition, it is not clear at what stage previous samples have been collected.

- Data collection stopping rule: We will stop collecting observational data and blood samples (for hormones and genetics) after we have completed research on the three populations included in this investigation (likely in 2022).
  What does it mean? When funding runs out? Is the aim to catch and sample every possible bird in the three populations? Knowing the type of analysis that will be performed, the number of response and explanatory variables etc. what would be a decent sample size to detect let's say a moderate effect with 80% chances (power analysis)? Is there hope to collect enough male carers during these years (i.e. is it possible to get a rough estimation of the expected number of male carer found each year based on the number of times male care was previously observed in relation to observation effort)? I do think those males should be sampled even if at the end the sample size is too small for robust statistics. Because they occur so rarely, every individual counts. However, it is good to get a realistic idea of the time it would take to answer the initial questions of the project with decent statistical power.

- Blinding of conditions during analysis: The researchers analyzing the DNA samples will be blind to the sex and identity of the grackle from which the sample came.
  State the same about people analysing hormonal contents.

- Independent variables: Perhaps call them "explanatory variables". Since this is not a controlled experiment, they are certainly not all independent.

- Note: if scaled mass index and hematocrit are correlated, we will choose only one of these variables to include in the analyses.
  Why? Why would that be a problem? This should not be a problem at all unless they are extremely correlated. (Morrissey & Ruxton 2018).

- H1: hormones (P1-P3)

- Analysis: We will run a Generalized Linear Model (GLM; glm function, stats package)
  This is a general linear model (lm) since the distribution is Gaussian. Will there be no repetition per individual? Getting data for several years gives the possibility to look at intra-individual variation depending on whether individuals hold territories or not and whether they provision food or not.

- “will determine whether an independent variable had an effect or not“
  Well, you will determine whether you detect an effect or not, but one should not conclude that there is no effect when one does not detect a statistically clear effect (i.e. one should not accept the null hypothesis).

- H4: mating strategy
  A GLM was conducted as in H1 above except this GLM has a Poisson distribution with a log link.
  Here it is indeed a generalised linear model I agree. However, isn’t the response variable binary (male care yes/no)? If so a Bernoulli distribution should be used (binomial) with e.g. a logit (or probit) link function.

References:

Alonso-Alvarez, C., & Velando, A. (2012). Benefits and costs of parental care. The evolution of parental care, 40-61.
Gustavo S. Requena, Glauco Machado, Effects of egg attendance on male mating success in a harvestman with exclusive paternal care, Behavioral Ecology, Volume 26, Issue 3, May-June 2015, Pages 926–935, https://doi.org/10.1093/beheco/arv035
Hadfield, J. (2012). The quantitative genetic theory of parental effects. The evolution of parental care, 267-284.
Lessells, C. K. M. (2006). The evolutionary outcome of sexual conflict. Philosophical Transactions of the Royal Society B: Biological Sciences, 361(1466), 301-317.
Morrissey, M. B., & Ruxton, G. D. (2018). Multiple regression is not multiple regressions: the meaning of multiple regression and the non-problem of collinearity.
Smiseth, P.T. What is Parental Care? (2012) The evolution of parental care

Reviewed by André C Ferreira, 2020-03-27 14:57

This project aims at studying the rare occurrence of male parental care behaviour of great-tailed grackles. Few observations of this behaviour have been collected and described, as well as several potential hypotheses for the occurrence of this behaviour have been proposed by the authors. I personally value the efforts that the authors made to describe the direct observations and to show videos and pictures of this rare behaviour. Such detailed descriptions help the readers to better understand the study species and the conditions in which these observations are made. I also value the risk that the authors are taking in studying a rarely observed behaviour. While I consider studying rare behaviours of extreme importance in the field of behaviour and evolution, I also know from experience that it can be extremely hard and frustrating to collect a large enough sample. This is in fact my biggest concern about this pre-study.

The authors have suggested testing several hypotheses that are grouped in four distinct topics. Being a rare behaviour, it is likely that by the end of 2022 the number of male great-tailed grackles observed feeding fledglings will be low (maybe less than 20?). Testing many hypotheses on such a small sample will most likely lead to type I errors. For this pre-study, it would be important to add an estimation of how many males the authors expect to follow or the number of males in the populations. This would help the readers to better understand the feasibility of study. I also suggest adding the description of the method for observing parental care behaviour. From what I understood (without much information) it seems to be direct focal observations. I strongly suggest the authors to study the possibility of using a more automated methodology, such as artificial feeding stations together with video cameras and/or RFID technology, in order to increase the sampling effort and the probability of observing the birds feeding the fledglings. RFID technology cannot directly determine if the males are feeding the fledglings but it could be used to see how associated the males are to the fledglings and indirectly infer about their parental care. Maximizing the sample effort is crucial for the feasibility of this study.

Another concern about this study is the lack of emphasis on understanding which conditions promote this rare behaviour. It would be important to determine for example, if the density of fledglings in the population

increases the proportion of males expressing parental care. Or if this behaviour is related with the food availability, as in harsh environments males could be more reluctant to feed fledglings than in places where food is abundant and therefore the cost of feeding fledglings is low. Understanding the conditions that promote the behaviour would not only be of great interest per se, but also help testing the proposed hypotheses. For example in “hypothesis 1” it is suggested to test if the males that express parental care behaviour have an hormone profile more similar to females than to other males that do not express parental care. Failing to reject the null hypothesis might be only a consequence of having males in the non-parental care group that did not express the behaviour simply because there were no (or few) fledglings available to feed and hence little opportunity to express the behaviour. The authors should either only compare males with territories with similar fledgling density or test if fledgling density is related with the proportion of males expressing parental care.

The hypotheses and predictions generally sound appropriate. I think that hypotheses 1 and 2 are crucial to understand the behaviour. Regarding hypothesis 3 I consider to be more important to compare the differences between the different populations in terms of population density, sex ratio, food availability and reproductive success (i.e. number of fledglings available to feed) than looking at the problem in terms of local adaptation. Recently established population probably already differ from well-established populations in one (or more) of the variables that I mentioned and therefore it would be more informative to know which variable(s) is(are) correlated with this behaviour, rather than comparing recently established populations vs older populations and speculate about what promotes the behaviour. Furthermore this hypothesis seems to arise from the fact that in spite of a greater coverage of sample effort of observations during the breeding season has been done in in well-established populations, most of the anecdotes of male parental care reported so far come from recently established populations. I warn the authors that this interpretation might be misleading as this behaviour seems to occur mostly when the chicks fledge which means that only studies that specifically look at the social interactions of the fledglings with the adults would be able to observe this rare behaviour and this might have not been the case in the previous studies that are cited.

I think this pre-study (and any manuscript that will result from this pre-study) would greatly benefit from changing the title. As it is now, only researchers that are familiarized with the reproductive system of the great-tailed grackles can immediately recognize the importance of studying male parental care in this species from reading the title. All other researchers will not understand the impact that this pre-study might have. I advise the authors to include in the title a reference to the fact that they are trying to study an atypical or rare form (or manifestation) of parental care. This way it would immediately catch the attention of researchers studying parental care in different taxa and not only researchers working with birds or that know the great-tailed grackles system.

Finally, I really value the fact that the authors did not look at this rare behaviour as an accidental or “misdirected” parental care and decided to study it instead. For example, some species that are not considered as cooperative breeders have been reported to have helpers in some populations (e.g. appendix 1.3 in Chapter 1, Ecology and Evolution of Cooperative Breeding in Birds. Koenig, W. D., & Dickinson, J. L. 2004, for a list of species that sometimes have helpers but are not considered cooperative breeders). Nevertheless most researchers studying cooperative breeding species promptly classify these observations as accidental or misdirected parental care just because of being of rare occurrence. I believe that there is much to learn from rarely observed behaviours and therefore I encourage the authors to pursue with this study as I believe that will be of interest to a broad audience.

Author's reply:

Dear Dr.’s Marie-Jeanne Holveck, Matthieu Paquet, and André Ferreira, We greatly appreciate the time you have taken to give us such useful feedback! We are very thankful for your willingness to participate in the peer review of preregistrations, and we are grateful to have the opportunity to revise and resubmit. We revised our preregistration and associated files at http://corinalogan.com/Preregistrations/gmalecare.html, and we responded to your comments below. Note that the version-tracked version of this preregistration is in rmarkdown at GitHub:

https://github.com/corinalogan/grackles/blob/master/Files/Preregistrations/gmalecare.Rmd. In case you want to see the history of track changes for the preregistration at GitHub, click the previous link and then click the “History” button on the right near the top. From there, you can scroll through our comments on what was changed for each save event and, if you want to see exactly what was changed, click on the text that describes the change and it will show you the text that was replaced (in red) next to the new text (in green). We think the revised version is much improved due to your generous feedback! All our best, Melissa, Maggie, Dieter, Kelsey, Luisa, Angela, Aaron, Carol, and Corina

Repeated parental care by adult male great-tailed grackles and its association with hormones, fitness, specific populations, and mating strategies Folsom MA, MacPherson M, Lukas D, McCune KB, Bergeron L, Bond A, Blackwell A, Rowney C, Logan CJ http://corinalogan.com/Preregistrations/gmalecare.html version 2.0 Submitted by Corina Logan 2019-12-05 17:38 Abstract This is a PREREGISTRATION submitted for pre-study peer review. Our planned data collection START DATE is May 2020, therefore it would be ideal if the peer review process could be completed before then. Abstract: Great-tailed grackles (Quiscalus mexicanus) are known to have a mating system where females are the sole builders of the nest and caretakers of the young (Johnson et al. (2000)). However, there is one report from Selander (1970) in which an unmarked male great-tailed grackle provided parental care to two juveniles in Austin, Texas. Here, we report repeated parental care events by multiple individually marked adult males in a population of great-tailed grackles in Tempe, Arizona. We also report male parental care by unmarked adult males in Tempe, Arizona; Santa Barbara, California; and San Clemente, California. We then propose to test hypotheses to determine whether male parental care 1) is associated with hormone profiles (testosterone, estrogen, prolactin, progesterone) potentially because hormones mediate investment in competition and care behavior; 2) increases direct fitness; 3) is a local adaptation, unique to specific populations where it has evolved; and 4) is a mating strategy. Keywords: birds, great-tailed grackle, parental care, male care, hormones, fitness, adaptation, mating strategy Round #1

Your decision by Marie-Jeanne Holveck, 2020-03-30 12:15 Manuscript: http://corinalogan.com/Preregistrations/gmalecare.html Ask for a revision

Dear Folsom MA, MacPherson M, Lukas D, McCune KB, Bergeron L, Bond A, Blackwell A, Rowney C, Logan CJ,

COMMENT 1: Your preprint entitled "Repeated parental care by adult male great-tailed grackles and its association with hormones, fitness, specific populations, and mating strategies" has now been reviewed and the reviewers’ comments are appended below. You will see that, while they find your work of interest, they have raised points that need to be addressed by a revision. Both reviewers and I strongly value your effort to program a study, in natural field conditions, of a rare, yet likely evolutionary-important behaviour, namely parental care by males of the great-tailed grackle. Yet, we all also recognised that it is a risky endeavour, and as a consequence, both reviewers wonder about your ability to reach a sufficient sample size to statistically test (all) hypotheses and predictions with enough confidence (e.g. risk of type I errors, also known as false positives). For instances, Reviewer 1 thus suggests running prior power analyses to clarify this issue; Reviewer 2 suggests some automated methodologies to boost the sampling effort. Both reviewers also have a series of additional relevant comments, notably regarding the null hypotheses’ identity, the rationale behind some of the predictions and behind the hypothesis 3, the age of the birds, the stopping rules for data collection, the statistical analyses, the conditions responsible for the occurrence of male parental care, the catchiness of the title, etc. (please see the detailed reviews). I believe that addressing their comments properly will highly improve your project. We therefore invite you to revise and resubmit your preprint, taking into account all of the points raised. We look forward to hearing from you soon. Best regards, Marie-Jeanne Holveck, Recommender for PCI Ecology

RESPONSE 1: Thank you so much for your feedback and for the opportunity to revise and resubmit! We think your feedback and the reviewer comments were relevant and we look forward to both addressing them below and in the revised preregistration.

Reviews Reviewed by Matthieu Paquet, 2020-02-04 11:14

COMMENT 2: The present pre-print aims to better understand the appearance and the fitness consequences of male post-hatching (post fledgling?) care, a very rarely observed behaviour in the study species, the great-tailed grackle (Quiscalus mexicanus). One important aim of the study is to look at whether such behaviour correlates with the circulating levels of several relevant hormones. I thank the authors for taking the time to pre-register their study plan. I found the topic interesting and stimulating. It is a particularly risky project since the study behaviour seems very rare (hence decent sample sizes may be hard to reach) but obviously worth investigating. I have some comments and suggestions that I hope may help the authors to get the best out of their project. (Feel free to contact me if some of the referred literature is not accessible to the authors)

RESPONSE 2: We really appreciate your support! And thank you very much for the offer to contact you. Your comments were really clear (and so helpful, thank you very much!) and your references were sufficient for us to find the additional information so we think we should be ok to revise with just this information, but please do let us know if we misinterpreted anything.

COMMENT 3: A first general comment: It would be great to give some rough estimation of prior probability of the different predictions based on current knowledge of the study species and on what is known in other species. I am not saying that these should be all ranked according to a prior probability nor given any number. But it is important to state whether the different predictions seem likely or not, first to judge whether they seem worth testing or not and second to later on have an honest discussion on whether the results were expected or not and how likely they are to be false positive. (One can always write all possible predictions and then later on state that the results were predicted). Finding statistically "significant" results when testing very unlikely hypothesis is highly likely to be a false positive http://fpr-calc.ucl.ac.uk/.

RESPONSE 3: Given that our sample size is likely going to be really small, we won't be able to do statistical significance testing. Instead, our aim is to assess which of the different alternatives receives the most support with the data we will be able to collect. Therefore, we hope readers will understand the exploratory nature of this research. As such, we think it is important to list the various alternative predictions to ensure that we are able to correctly assess whether a given prediction might have some support (rather than setting up a false dichotomy between our favorite prediction and an alternative that is not likely to be supported).

Additionally, we have run into a lack of information about male parental care in non-monogamous bird species with regard to hormone profiles, so our predictions are all very speculative, and, therefore, all perhaps equally plausible/implausible. We run into the same problem with the relatedness and mating strategy hypotheses: we really don't have enough information to be able to suggest whether some predictions are more likely than others, so we would prefer to refrain from speculating.

To make the exploratory nature of our study clearer, we added:

Hypotheses: "...We note that this study will be more of an exploratory nature because of the likely very small sample size of males that provide parental care in this species. Our aim is to assess which of the predictions receive the most support based on the amount of data we are able to collect."

COMMENT 4: A second general comment: Is there any information on the age of the birds? If so, it would be interesting to know if males providing care are essentially young or old individuals. This would give important information on the understanding of the system, notably depending on whether these individuals happen to be genetic parents, social parents, or alloparents (i.e. helpers). I am not saying that age should be included in the statistical models (there will certainly not be enough power to detect any age effect on the top of the rest and it is not a cofounding factor), but that it could be looked at in a more descriptive way.

RESPONSE 4: We agree that the age of the adult males providing care would be an interesting variable to include in our study. Unfortunately, we are unable to age adult grackles so we will not be able to include this variable. We summarize here our explanation about why we are unable to obtain this data from our response in a peer review process for a separate preregistration (http://corinalogan.com/Preregistrations/gspaceuse.html): We agree that knowing the age of the male carers would give us additional important information. While Pyle (1987) suggests that it is possible to distinguish

second year grackles from first year grackles, in practice we do not observe consistent differences that align with these descriptions. This is likely due to the large population differences in biometrics and other morphometrics that we are finding. Therefore, unfortunately, we are not able to determine the age of an adult grackle after it's first year so we cannot add the age of these male carers to the data set (other than first year or adult).

Pyle, P. (1987). Identification guide to North American passerines: a compendium of information on identifying, ageing, and sexing passerines in the hand. Slate Creek Press.

Please find below some more specific comments/suggestions:

COMMENT 5: Prediction 2 (additive): If high T (thus lower E, prolactin, and progesterone) Why lower prolactine in this case? Since prolactin is linked with parental care one would assume that males that provide care have higher T than males that do not have territories but also higher prolactine than males without territories and males with territories but that don't provide care?

RESPONSE 5: This is a good point and we have now refined the prediction accordingly: Hypotheses > P2 (additive): "If high T is necessary to become a territory holder, and only territory holders provide care, then we expect T in caring males to be the same as in males who hold a territory but do not provide care, and higher than in non-territory holding males. This is because, in males, T is positively correlated with breeding success (@garamszegi2005testosterone), and is higher in breeding males than in helper males (in cooperatively breeding species; @pikus2018testosterone). If parenting is additive and dependent on condition, we expect these same high T males to also have higher prolactin and estrogen, and to engage in more parenting behaviors."

COMMENT 6: Prediction 3 (independent): Hormone levels do not systematically differ across individuals, regardless of sex or caring behavior, indicating that these hormones are not involved in parental care. Is this really predicted knowing prior knowledge on the role of these hormones in parental care in birds? It is important to test for a priori likely hypotheses in order to minimise the chance of false discoveries. This is rather the null hypothesis and I doubt it will be tested. I rather guess that the other hypotheses will be tested against this one.

RESPONSE 6: Thanks for catching this! We realize what we wrote for this didn't really reflect our intention. We have refined the prediction accordingly:

Hypotheses > P3 (independent): "T will be associated with territorial behavior, and prolactin and E will be associated with parenting behaviors, but these will vary independently (no positive or negative correlation)."

COMMENT 7: H2: Perhaps name it "relatedness" as local adaptation for hypothesis 3 sounds quite genetic as well.

RESPONSE 7: Good point! We changed "genetics" to "relatedness" in the headers throughout the preregistration.

COMMENT 8: Prediction 4 (direct fitness): Males who provide parental care increase their direct fitness by caring for their own offspring. The more care offspring receive, the more likely they are to survive, therefore any male care will have a direct fitness benefit. This prediction makes sense but note that this is not affecting their direct fitness. This is affecting their offspring's fitness (and therefore affecting male inclusive fitness). Testing for direct fitness benefits would mean testing whether caring males survive better or increase their future number of offspring (egg laid), for example if providing care increases their attractiveness towards females, which is a possibility (e.g. Requena, & Machado 2015 and cf alternative 4). Of course it is always hard to draw a line between parents and offspring fitness in species providing post egg-laying care, but this is where the line is drawn in quantitative genetics anyway (e.g. Hadfield 2012, Smiseth 2012). To make is less problematic it would for sure be fine to just state that males increase their offspring fitness (instead of mentioning their direct fitness).

RESPONSE 8: We realize that these terms (direct and indirect fitness) have different meanings in different fields. To avoid any confusion, we now omitted "direct" and "indirect" and rephrased to refer to the specific benefit. We made the following revisions:

Hypotheses > Hypothesis 2: "Paternal care increases the number of offspring that survive to independence"

Hypotheses > Hypothesis 2 > Prediction 4 (own offspring): "Males who provide parental care do so for their own offspring. The more care offspring receive, the more likely they are to survive, therefore any male care will have a direct influence on their offspring's fitness"

Hypotheses > Hypothesis 2 > Prediction 4, alternative 2 (offspring of relatives): "If the males providing parental care are not the parents of the offspring cared for, we predict that these males are caring for the offspring of close relatives. For example, Nacho and Taquito (males who hold neighboring territories) could be Chile's brothers and/or Crema (observed near Chile and the two fledglings) could be his sister. Therefore, if Pepino is an offspring of one of these nearby individuals, Chile would be helping his relative."

COMMENT 9: Prediction 4, alternative 4 (signal): If the caring male is not the father of the offspring he is caring for, the caring behavior might be a signal to the offspring's mother, which then increases his chance of fathering her next offspring. Note that this may be true even if he IS the father. Females could still use this cue as a signal affecting whether or not they will breed with the same male again.

RESPONSE 9: Good point! We updated this prediction accordingly:

Hypotheses > Hypothesis 2 > Prediction 4, alternative 4 (signal): "The caring behavior might be a signal to the offspring's mother, which then increases the caring male's chance of fathering her next offspring, independently of whether he is or is not the father of the offspring he is currently caring for."

COMMENT 10: Prediction5: Perhaps male parental care is more prevalent near the edge of their range because the costs of raising young might be higher (e.g., establishing in a new location, exposure to new predators, less protection due to a lower conspecific population density) and require a higher investment from parents. This sounds counter-intuitive. Higher cost of care should reduce the probability for parental care to occur (if benefits stay constant). What matters is the cost-benefit balance (See e.g. Lessells 2006, Alonso-Alvarez, C., & Velando, A. 2012).

RESPONSE 10: We see your point that we needed to separate out the costs and benefits from the perspective of the males and the offspring. Therefore, we modified the original Prediction 5, and we added a new alternative (P5 alternative 1):

Hypothesis 3 > Prediction 5 (relative benefits of paternal care are higher at the range edge because offspring benefit more from care or because males have lower costs): "...Perhaps male parental care is more prevalent near the edge of their range because the offspring are more likely to benefit from receiving care. For example, offspring might need more protection from predators, they may need to learn more about foraging options, or they might be in a place with higher seasonality where the time for providing care is shorter. For males, the relative benefits of providing this extra investment might outweigh their potential costs."

Hypothesis 3 > Prediction 5, alternative 1 (relative benefits of paternal care are lower at the range edge because offspring benefit less from care or because males have higher costs): "Perhaps male parental care is less prevalent near the edge of their range because the costs of raising offspring are higher. For example, costs to males of providing care in edge populations might be higher because they might be more exposed to new predators, it might be more difficult to find food in a new location, there might be a female-skewed sex ratio (Sevchik et al. 2019) such that males will lose out on additional mating opportunities if they provide care, or there might be less protection due to a lower conspecific population density. For males, the relative costs of providing this extra investment might not outweigh the potential benefits."

COMMENT 11: Prediction 10 (no relationship): Male parental care is not related to physical condition or whether they hold a territory. As for prediction 4, is that very likely that parental care is not related to physical condition or territory holding? These 2 factors surely affect the cost of care so they are predicted to

affect the occurrence of parental care. This is rather a null hypothesis. When will blood sampling be conducted? Will they be sampled during both breeding and non breeding season? It would be great to have also blood samples of those food provisioning males when they are not breeding (and some non-caregiving males and females as well as control) in order to better understand the link between food provisioning and hormonal levels, also using intra-individual comparisons (do male care takers have high levels anyway or does their level vary depending whether they provide food or not?). Since we can expect a quite small number of male providing care, it would be a very precious additional information to get from them. In addition, it is not clear at what stage previous samples have been collected.

RESPONSE 11: You are correct in that Prediction 10 is the null hypothesis. Regarding the likelihood of Prediction 10 relative to other predictions in Hypothesis 4 (about territory holding and condition), because we have so few data points at this stage, it is really difficult to speculate about what will and will not occur, so we wanted to present a range of possible options and their potential explanations. This kind of question comes up a lot when we submit preregistrations for pre-study peer review and we like to clarify why we think it is important to list our various predictions in advance. For this, we quote our response to a reviewer in the peer review process for a different preregistration at PCI Ecology (Mendez 2019): "For each hypothesis, there are a number of results that could occur (e.g., positive, negative, or no correlations) and we wanted to make a priori predictions about how we would interpret every potential result from a given hypothesis. This prevents us from HARKing (Hypothesizing After Results are Known; see Kerr 1998), which could occur if we get a result that we weren't expecting. In this case, we could then make up a post hoc story about why that result might have occurred. By a priori accounting for as many variations of the results that we can think of, it places our focus on being predictive in advance, which allows us to test these predictions in this study (see Nosek et al. 2019). If we didn't list the alternatives at the pre-data collection stage, and we ended up encountering a result that was not in our predictions, we would be providing an interpretation post hoc, which would require us to conduct a new study to determine whether that prediction was supported. Another advantage to listing multiple alternatives in advance and having automated version tracking at GitHub with time and date stamps and track changes for all edits to the document is that readers can verify for themselves whether we were HARKing or not. Listing all potential predictions in advance allows us to explore the whole logical space that we are working in, rather than just describing one outcome possibility." We agree that intra-individual measurements would add so much to this dataset. Unfortunately, we are not able to recapture individual great-tailed grackles (see Response 16 for a more detailed explanation) so we only have one blood sample per male. We catch both sexes in the breeding and non-breeding seasons and we have so far always had individuals in the sample who are non-caring males as well as caring and non-caring females. We realize this might not have been clear in the preregistration so we fixed the code for Table 2 where this is summarized (only for the breeding season) so now the column titles are not cut off and you can see the sample sizes for each category so far. We decided to only analyze data from blood samples we collect during the breeding season to reduce the potential confound that breeding and non-breeding season hormonal profiles differ from each other. We see though that this wasn't very clear from our write up so we have now clarified these points:

Methods > Planned sample: "...Note that we are unable to recapture individuals, therefore there is only one blood sample per individual and it can be from the breeding or non-breeding season. Although we collect data on this project in both the breeding and non-breeding seasons, we will limit our analyses in this study to only the breeding season data to reduce potential confounds between breeding and non-breeding season hormone and behavioral profiles."

Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., ... & Vazire, S. (2019). Preregistration Is Hard, And Worthwhile. Trends in cognitive sciences, 23(10), 815-818.

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. Personality and Social Psychology Review, 2(3), 196-217.

Marcos Mendez (2019) Are condition indices positively related to each other and to fitness?: a test with grackles. Peer Community in Ecology, 100035. 10.24072/pci.evolbiol.100035

COMMENT 12: Data collection stopping rule: We will stop collecting observational data and blood samples (for hormones and genetics) after we have completed research on the three populations included in this investigation (likely in 2022). What does it mean? When funding runs out? Is the aim to catch and sample every possible bird in the three populations? Knowing the type of analysis that will be performed, the number of response and explanatory variables etc. what would be a decent sample size to detect let's say a moderate effect with 80% chances (power analysis)? Is there hope to collect enough male carers during these years (i.e. is it possible to get a rough estimation of the expected number of male carer found each year based on the number of times male care was previously observed in relation to observation effort)? I do think those males should be sampled even if at the end the sample size is too small for robust statistics. Because they occur so rarely, every individual counts. However, it is good to get a realistic idea of the time it would take to answer the initial questions of the project with decent statistical power.

RESPONSE 12: We updated the text to clarify that we will stop collecting data when the current funding period ends:

Methods > Data collection stopping rule: "We will stop collecting observational data and blood samples (for hormones and genetics) after we have completed research on the three populations included in this investigation (likely in 2022), which coincides with the period in which we have funding for this project."

Good point that we needed to provide more details about our plans for collecting data at each of the field sites. We aim to catch and color-mark as many individuals as possible in each population so that we can collect more detailed data on their social interactions and learning. To address this, we added:

Methods > Planned sample:

"Great-tailed grackles (n > 200) will be caught in the wild at three field sites across their geographic range: the center of their original range (at a site to be determined in Central America), the middle of the northward expanding edge (Tempe, Arizona USA), and near the northern expanding edge (Woodland, California USA). Individuals will be identified using colored leg bands in unique combinations, their data collected (blood, feathers, and biometrics), and then they will be released back to the wild where we will collect data on their behavior." We added power analyses to determine what our ability to detect effects is based on what we think will probably be a very small sample size (we estimated it at 9 grackles based on what we've found in Arizona; see our Response 19 for more details on our expectations about the sample size).

Analysis Plan > Ability to detect actual effects:

"To understand what effect sizes we will be able to detect given our sample size limitations and the number of explanatory variables, we used G*Power (v.3.1, @faul2007g, @faul2009statistical) to conduct power analyses based on confidence intervals. G*Power uses pre-set drop down menus and we chose the options that were as close to our analysis methods as possible (listed in each analysis below). We realize that these power analyses are not fully aligned with our study design, however we are unaware of better options at this time. Additionally, it is difficult to run power analyses because it is unclear what kinds of effect sizes we should expect due to the lack of data on this species for these measures."

And we added the results of the power analyses to two analyses:

Analysis Plan > H1 hormones (P1-P3)

Analysis Plan > H4 mating strategy

The results show that we would have the ability to detect only very large effect sizes. This isn't a surprising result, but we think it will be worth it to give the study a try and see how far we can get.

COMMENT 13: Blinding of conditions during analysis: The researchers analyzing the DNA samples will be blind to the sex and identity of the grackle from which the sample came. State the same about people analysing hormonal contents.

RESPONSE 13: Yes, good point, we are able to do this for the hormone samples as well so we changed the preregistration accordingly: Methods > Blinding of conditions during analysis: The researchers analyzing the DNA and hormone samples will be blind to the sex of the grackle from which the sample came (the bird's ID is written on the tube, however, the people performing the lab work will not have field experience with the grackles, therefore they will not know the grackle's sex or other information about that individual based on their ID).

COMMENT 14: Independent variables: Perhaps call them “explanatory variables”. Since this is not a controlled experiment, they are certainly not all independent.

RESPONSE 14: Of course! We changed “independent variables” to “explanatory variables”. We also changed “dependent variables” to “response variables”.

COMMENT 15: Note: if scaled mass index and hematocrit are correlated, we will choose only one of these variables to include in the analyses. Why? Why would that be a problem? This should not be a problem at all unless they are extremely correlated. (Morrissey & Ruxton 2018).

RESPONSE 15: We wanted to reduce the number of explanatory variables in the model as much as possible because of the very likely possibility that we will have an extremely small sample size. Because we are using both of these measures as a proxy for condition, we thought that, rather than doubling up, we would prefer to just use one, which gives us more power for the analysis.

COMMENT 16: H1: hormones (P1-P3)

a) Analysis: We will run a Generalized Linear Model (GLM; glm function, stats package) This is a general linear model (lm) since the distribution is Gaussian. Will there be no repetition per individual? Getting data for several years gives the possibility to look at intra-individual variation depending on whether individuals hold territories or not and whether they provision food or not.

b) “will determine whether an independent variable had an effect or not“ Well, you will determine whether you detect an effect or not, but one should not conclude that there is no effect when one does not detect a statistically clear effect (i.e. one should not accept the null hypothesis).

RESPONSE 16:

a) Thank you for catching that! We changed Analysis Plan > H1 P1-P3 to: General Linear Model. We only have repeated measures for one male (Chile) because it is almost impossible to recapture the same individuals. So, unfortunately, we will have to stick with GLMs rather than using GLMMs and including individual as a random variable. Based on your comment, we realized it might not be clear from our preregistration that we aren’t spending several years at each field site because we only have enough resources to work at one field site at a time. This has resulted in us spending 3 years in Arizona, and we plan to spend one year each at the next two field sites. To address this, we added the years we plan to be at each field site to Methods > Planned sample. Spending only one year at each of the other sites reduces our ability to look at intra-individual variation, which would be limited anyway because of our lack of ability to recapture individuals.

b) We realize that with our sample size we will only have the power to detect very large effects. Accordingly, we revised the preregistration to say: Analysis Plan > H1 P1-P3 and H4 mating strategy: “We will determine whether an explanatory variable had a detectable effect using the Estimate in the full model: if the 89% confidence intervals do not overlap zero, we will consider this a detectable effect (@statrethinkingbook).”

COMMENT 17: H4: mating strategy A GLM was conducted as in H1 above except this GLM has a Poisson distribution with a log link. Here it is indeed a generalised linear model I agree. However, isn’t the response variable binary (male care yes/no)? If so a Bernoulli distribution should be used (binomial) with e.g. a logit (or probit) link function.

RESPONSE 17: Nice catch! We changed Analysis Plan > H4 to: “We will run a Generalized Linear Model (GLM; glm function, stats package) with a binomial distribution and logit link. ”

Reviewed by André C Ferreira, 2020-03-27 14:57

COMMENT 18: This project aims at studying the rare occurrence of male parental care behaviour of great-tailed grackles. Few observations of this behaviour have been collected and described, as well as several potential hypotheses for the occurrence of this behaviour have been proposed by the authors. I personally value the efforts that the authors made to describe the direct observations and to show videos and pictures of this rare behaviour. Such detailed descriptions help the readers to better understand the study species and the conditions in which these observations are made. I also value the risk that the authors are taking in studying a rarely observed behaviour. While I consider studying rare behaviours of extreme importance in the field of behaviour and evolution, I also know from experience that it can be extremely hard and frustrating to collect a large enough sample. This is in fact my biggest concern about this pre-study.

RESPONSE 18: We really appreciate your positive feedback and for sharing with us what you have learned about trying to collect data on rare behaviors!

COMMENT 19: The authors have suggested testing several hypotheses that are grouped in four distinct topics. Being a rare behaviour, it is likely that by the end of 2022 the number of male great-tailed grackles observed feeding fledglings will be low (maybe less than 20?). Testing many hypotheses on such a small sample will most likely lead to type I errors. For this pre-study, it would be important to add an estimation of how many males the authors expect to follow or the number of males in the populations. This would help the readers to better understand the feasibility of study. I also suggest adding the description of the method for observing parental care behaviour. From what I understood (without much information) it seems to be direct focal observations. I strongly suggest the authors to study the possibility of using a more automated methodology, such as artificial feeding stations together with video cameras and/or RFID technology, in order to increase the sampling effort and the probability of observing the birds feeding the fledglings. RFID technology cannot directly determine if the males are feeding the fledglings but it could be used to see how associated the males are to the fledglings and indirectly infer about their parental care. Maximizing the sample effort is crucial for the feasibility of this study.

RESPONSE 19: We think it is highly likely that the number of males observed to engage in male parental care will be <20 by the end of the study period. We agree that if the sample size is really small, we might not be able to detect whether some/any of our predictions are supported. If this is the case, we would hope that our hypotheses, predictions, and whatever data we are able to collect, will give an indication of how prevalent this behavior is and help drive future research on this topic that might be able to finish testing the hypotheses we outline here (see our Response 3 for a more detailed explanation regarding our aim, which is not to conduct statistical significance testing, but to assess the potential support for the various predictions). We are unsure of whether male parental care is present in either of the next two populations that we will establish field sites at. Logan and Lukas scouted the northern California field site in the summer of 2019 and spoke with birders there who had reported the first observations of this species in the area. When asked whether they had observed any male parental care, they said they weren't looking for it so they didn't have a memory of having seen it. We have not yet scouted the Central American field site so we are unsure of whether anyone there has noticed male parental care in the grackles in that population. Therefore, our most conservative estimation of how many more males we are likely to observe conducting male parental care is zero. A more optimistic estimate would assume that we might encounter the same number of males engaging in parental care at the other two populations as we have observed in Arizona, in which case the estimate would be at least three per population (total n=9). Given that we have observed two sightings of male parental care in southern California, we think it is likely that we will observe this behavior at the northern California site. We updated the preregistration as follows: Methods > Sample size rationale:

"We are unsure of how many males will provide male parental care at the three populations under study. However, given that we have observed at least three males engaging in parental care in Arizona, and observations of this behavior in two locations in southern California, we think it is likely that we will observe at least one male providing care at the northern California field site, but unlikely that we will observe large numbers because this behavior has been so rarely noted. We are not sure what to expect from the Central

American field site because no reports of male parental care have come from this region. Therefore, it is likely that our sample size will be <15 males who provide parental care." Great point that we needed to expand more on the behavioral observations we collect as parts of other preregistrations. We added details to this section in the preregistration to clarify: Methods > Collecting behavioral observations: "As part of separate preregistrations, we conduct two to three types of systematic field observational data on the grackles at each study site. These data collection protocols provide the opportunity to discover and record male parental care in a more systematic way rather than simply relying on chance anecdotes. We will continue collecting focal follows and space use data, as well as anecdotes, during the breeding seasons at the next two study populations.

1) Nest checks are conducted at the Arizona study site. We find nests by following radio-tagged grackles and non-radio tagged grackles to their nest sites and by searching for the nests of color-marked grackles in our study area. We record the females and males that are present in the nesting area, identify the nest stage (e.g., building, incubating, nestling), and take a GPS point of the nest location. See the full data collection protocol.

2) Focal follows to collect data on 19 social and foraging behaviors, which includes food sharing behavioral data that the carers of offspring engage in. We aim to conduct a minimum of four 10-min focal follows per bird across the breeding and non-breeding seasons. See the ethogram of behaviors we collect data on.

3) Track space use of the color-marked grackles in our study. We obtain 20+ GPS points on an estimated ~60 radio-tagged and ~60 non-radio-tagged grackles during the breeding and non-breeding seasons (~20 grackles of each type at each of the three sites). After finding an individual to track, the grackle is followed for up to 90 min, collecting a GPS point every minute as well as recording the behavior the grackle was engaged in that minute. See the full data collection protocol."

In terms of automating this type of data collection, for a few years we have been discussing the possibility of using ICARUS satellite tags (https://www.icarus.mpg.de/en) to track the movements of the individuals in our study. The tags are still too heavy to use on grackles (according to the USGS Bird Banding Laboratory restrictions), however the ICARUS team is working on making lighter models, which they hope to begin testing this year. Thus, we hope that we might be able to implement this solution sometime near the end of this 5 year project. In terms of implementing other types of automated data collection methods to aid in the detection of male parental care, the field team is currently working at maximum capacity to collect all of the other data we have committed to collecting on the project, therefore we do not have more resources to devote to the development of other potential avenues for automation. Particularly because methods such as video recording requires many staff hours to go through the videos and look for observations.

There are a few reasons we believe that our other data collection measures (above) are sufficient to detect male parental care based on what we have found in Arizona:

1) We attempt to find the territories of all breeding color-marked males in our study (so we can conduct the focal follows and GPS tracking).

2) Once found, the territories do not change locations, and many territories are adjacent to each other and observable at the same time.

3) Team members revisit the territories often as they collect data on this and nearby territories/nests/grackles.

4) The males we have observed providing care do so repeatedly (apparently daily), therefore, while this appears to be a behavior that is rare among males, once a male expresses it, he does so regularly, therefore making his behavior more detectable.

COMMENT 20: Another concern about this study is the lack of emphasis on understanding which conditions promote this rare behaviour. It would be important to determine for example, if the density of fledglings in the population increases the proportion of males expressing parental care. Or if this behaviour is related with

the food availability, as in harsh environments males could be more reluctant to feed fledglings than in places where food is abundant and therefore the cost of feeding fledglings is low. Understanding the conditions that promote the behaviour would not only be of great interest per se, but also help testing the proposed hypotheses. For example in "hypothesis 1" it is suggested to test if the males that express parental care behaviour have an hormone profile more similar to females than to other males that do not express parental care. Failing to reject the null hypothesis might be only a consequence of having males in the non-parental care group that did not express the behaviour simply because there were no (or few) fledglings available to feed and hence little opportunity to express the behaviour. The authors should either only compare males with territories with similar fledgling density or test if fledgling density is related with the proportion of males expressing parental care.

RESPONSE 20: We aren't sure whether we will have a large enough sample size to be able to understand the details about many of the conditions surrounding when male parental care occurs. Collecting data on the nestling density per territory is beyond the scope of our resources on this project: it takes up all of the field team's time to track the nests of the color-marked grackles while collecting the other breeding season data that we previously committed to (e.g., focal follows, GPS tracking). Therefore, we don't track non-color-marked grackle nests within the territories we monitor even though there are definitely nests from such grackles in the territories we track.

Regarding food availability, we collect data on a variety of aspects of the habitat and foods the grackles eat during focal follows that are part of a separate preregistration that has passed pre-study peer review at PCI Ecology (http://corinalogan.com/Preregistrations/g_flexforaging.html). The measurement of food availability came up in that peer review process as well (Astegiano & Sebastián González 2019) and our response is that it is not feasible for us to accurately measure: "[...] we aren't sure how we would quantify the amount of human food available. This is because we really only find the grackles in the middle of areas that are surrounded by human foods. For example, if we were to conduct a focal follow in a nearby park, there would be several trash cans scattered throughout the park and there might be a few people eating at the picnic tables. Given that we are not able to quantify how much non-human food is available at any given location (i.e., we can't see the worms under the grass or how many insects there are or how many fruits are in the area), then we aren't quite sure how to give this a quantitative measure. However, we will be able to quantify how frequently human foods are consumed because we can see what they eat and this is measured during the focal follows. We will also measure the substrate they were last observed on for each focal follow so this will give us an indication of whether that focal occurred in a grassy area (where they were more likely eating non-human foods) or on cement (where they would be more likely to be eating human foods)."

Unfortunately, we will not have detailed enough data to be able to investigate the interaction between fledgling density, food availability, and male parental care, but we hope that the discussion here, spurred by your insightful comment, will be useful for future studies that focus solely on male parental care (if we indeed find that it is present in multiple populations).

Julia Astegiano and Esther Sebastián González (2019) Understanding geographic range expansions in human-dominated landscapes: does behavioral flexibility modulate flexibility in foraging and social behavior?. Peer Community in Ecology, 100026. 10.24072/pci.ecology.100026

COMMENT 21: The hypotheses and predictions generally sound appropriate. I think that hypotheses 1 and 2 are crucial to understand the behaviour. Regarding hypothesis 3 I consider to be more important to compare the differences between the different populations in terms of population density, sex ratio, food availability and reproductive success (i.e. number of fledglings available to feed) than looking at the problem in terms of local adaptation. Recently established population probably already differ from well-established populations in one (or more) of the variables that I mentioned and therefore it would be more informative to know which variable(s) is(are) correlated with this behaviour, rather than comparing recently established populations vs older populations and speculate about what promotes the behaviour. Furthermore this hypothesis seems to arise from the fact that in spite of a greater coverage of sample effort of observations during the breeding season has been done in in well-established populations, most of the anecdotes of male parental care

reported so far come from recently established populations. I warn the authors that this interpretation might be misleading as this behaviour seems to occur mostly when the chicks fledge which means that only studies that specifically look at the social interactions of the fledglings with the adults would be able to observe this rare behaviour and this might have not been the case in the previous studies that are cited.

RESPONSE 21:

a) Because we are only comparing three populations, we cannot test specifically which variables might influence this behavior and that these influences are directly related to the population's location in the range. Instead, as a first step, we are investigating whether being in an edge population might be associated with prevalence of paternal care. We modified Hypothesis 3 Prediction 5 and added Prediction 5 alternative 1 to discuss the various factors that could be involved and how they might influence the potential costs and benefits of providing paternal care. The revision is as follows:

Hypothesis 3 > Prediction 5 (relative benefits of paternal care are higher at the range edge because offspring benefit more from care or because males have lower costs): "...Perhaps male parental care is more prevalent near the edge of their range because the offspring are more likely to benefit from receiving care. For example, offspring might need more protection from predators, they may need to learn more about foraging options, or they might be in a place with higher seasonality where the time for providing care is shorter. For males, the relative benefits of providing this extra investment might outweigh their potential costs."

Hypothesis 3 > Prediction 5, alternative 1 (relative benefits of paternal care are lower at the range edge because offspring benefit less from care or because males have higher costs): "Perhaps male parental care is less prevalent near the edge of their range because the costs of raising offspring are higher. For example, costs to males of providing care in edge populations might be higher because they might be more exposed to new predators, it might be more difficult to find food in a new location, there might be a female-skewed sex ratio such that males will lose out on additional mating opportunities if they provide care, or there might be less protection due to a lower conspecific population density. For males, the relative costs of providing this extra investment might not outweigh the potential benefits."

b) Regarding the sample effort during the breeding season invested by previous research, we agree that these studies might not have detected male parental care due to not being in the field at the right time. We made sure to note this in Prediction 5, alternative 2:

Hypothesis 3 > "Prediction 5, alternative 2 (common across range):** Male parental care occurs in populations across the range, regardless of establishment date. It might not have been considered a common behavior previously because of a lack of sampling effort during the breeding season at populations across their range."

COMMENT 22: I think this pre-study (and any manuscript that will result from this pre-study) would greatly benefit from changing the title. As it is now, only researchers that are familiarized with the reproductive system of the great-tailed grackles can immediately recognize the importance of studying male parental care in this species from reading the title. All other researchers will not understand the impact that this pre-study might have. I advise the authors to include in the title a reference to the fact that they are trying to study an atypical or rare form (or manifestation) of parental care. This way it would immediately catch the attention of researchers studying parental care in different taxa and not only researchers working with birds or that know the great-tailed grackles system.

RESPONSE 22: This is a really good point and we totally agree with you. We changed the title to: "Investigating the rare behavior of male parental care in great-tailed grackles"

COMMENT 23: Finally, I really value the fact that the authors did not look at this rare behaviour as an accidental or "misdirected" parental care and decided to study it instead. For example, some species that are not considered as cooperative breeders have been reported to have helpers in some populations (e.g. appendix 1.3 in Chapter 1, Ecology and Evolution of Cooperative Breeding in Birds. Koenig, W. D., & Dickinson, J. L. 2004, for a list of species that sometimes have helpers but are not considered cooperative

breeders). Nevertheless most researchers studying cooperative breeding species promptly classify these observations as accidental or misdirected parental care just because of being of rare occurrence. I believe that there is much to learn from rarely observed behaviours and therefore I encourage the authors to pursue with this study as I believe that will be of interest to a broad audience.

RESPONSE 23: Thank you so much for this (and for all of your other) feedback! One thing we have learned from studying two populations of great-tailed grackles so far is that their behavior can be surprisingly different across populations. When we saw male parental care, we immediately thought we had encountered some interesting population variation. Thank you so much for your support of this research!