



The role of behavior and habitat availability on species geographic expansion

[Esther Sebastián González](#) based on reviews by Pizza Ka Yee Chow, Caroline Marie Jeanne Yvonne Nieberding, Tim Parker and 1 anonymous reviewer

A recommendation of:

Open Access

Logan CJ, McCune KB, Breen A, Chen N, Lukas D. **Implementing a rapid geographic range expansion - the role of behavior and habitat changes (2020)**, *In Principle Recommendation*. *PCI Ecology*. <http://corinalogan.com/Preregistrations/gxpopbehaviorhabitat.html> version 06 Oct 2020

Submitted: 14 May 2020, Recommended: 05 October 2020

Cite this recommendation as:

Esther Sebastián González (2020) The role of behavior and habitat availability on species geographic expansion. *Peer Community in Ecology*, 100062. [10.24072/pci.ecology.100062](https://doi.org/10.24072/pci.ecology.100062)

Published: 6 October 2020

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

Understanding the relative importance of species-specific traits and environmental factors in modulating species distributions is an intriguing question in ecology [1]. Both behavioral flexibility (i.e., the ability to change the behavior in changing circumstances) and habitat availability are known to influence the ability of a species to expand its geographic range [2,3]. However, the role of each factor is context and species dependent and more information is needed to understand how these two factors interact.

In this pre-registration, Logan et al. [4] explain how they will use Great-tailed grackles (*Quiscalus mexicanus*), a species with a flexible behavior and a rapid geographic range expansion, to evaluate the relative role of habitat and behavior as drivers of the species' expansion [4]. The authors present very clear hypotheses, predicted results and also include alternative predictions. The rationales for all the hypotheses are clearly stated, and the methodology (data and analyses plans) are described with detail. The large amount of information already collected by the authors for the studied species during previous projects warrants the success of this study. It is also remarkable that the authors will make all their data available in a public repository, and that the pre-registration is already stored in GitHub, supporting open access and reproducible science.

I agree with the three reviewers of this pre-registration about its value and I think its quality has largely improved during the review process. Thus, I am happy to recommend it and I am looking forward to seeing the results.

References

- [1] Gaston KJ. 2003. The structure and dynamics of geographic ranges. Oxford series in Ecology and Evolution. Oxford University Press, New York.
- [2] Sol D, Lefebvre L. 2000. Behavioural flexibility predicts invasion success in birds introduced to new zealand. *Oikos*. 90(3): 599–605. doi: <https://doi.org/10.1034/j.1600-0706.2000.900317.x>
- [3] Hanski I, Gilpin M. 1991. Metapopulation dynamics: Brief history and conceptual domain. *Biological journal of the Linnean Society*. 42(1-2): 3–16. doi: <https://doi.org/10.1111/j.1095-8312.1991.tb00548.x>
- [4] Logan CJ, McCune KB, Breen A, Chen N, Lukas D. 2020. Implementing a rapid geographic range expansion - the role of behavior and habitat changes (<http://corinalogan.com/Preregistrations/gxpobbehaviorhabitat.html>) In principle acceptance by PCI Ecology of the version on 6 Oct 2020 <https://github.com/corinalogan/grackles/blob/master/Files/Preregistrations/gxpobbehaviorhabitat.Rmd>.

Reviewed by [Caroline Marie Jeanne Yvonne Nieberding](#), 2020-09-18 07:44

Review 2 of revised version of manuscript by Logan et al “Implementing a rapid geographic range expansion - the role of behavior and habitat changes” submitted to PCI Ecology

I would like to congratulate the authors for revising so thoroughly their manuscript, and I have no further comments or concerns after reading the responses to referees and revised manuscript. Except for one useful "detail" regarding comment and response 39: in fact, the actual original paper showing that learning is costly was by Mery and Kawecky in *Science* in 2005. Using experimental evolution on learning for oviposition in flies, they showed that learning induced a reduction in offspring production and in survival, as far as I remember.

Best wishes for collecting sufficient data, CN.

Reviewed by anonymous reviewer, 2020-10-02 15:56

Thank you for inviting me to review this revised pre-print manuscript. I think the authors did a great job in the revision – the investigation goal presented in the Introduction is clearer than the previous version, the information presented in the current revision is also consistent with the goal of the investigation. I understand that there are many areas to be explored in the study, but the authors have addressed my previous comments and concerns appropriately. Hence, there is no major issue to be raised.

Reviewed by [Tim Parker](#), 2020-09-18 18:34

This pre-registration draft is a plan for studying range expansion in great-tailed grackles. The authors present clear questions and predictions, and detailed analysis plans. This is an appropriate pre-registration.

This authors of this pre-registration have addressed nearly all the concerns I laid out in my review of the prior draft. I have only one recommendation for a change prior to archiving (see below)

However, as I stated in the original review, I wish to acknowledge that I lack expertise regarding some of the methods in this pre-registration, and therefore cannot attest to their sufficiency. In particular, I am unfamiliar with the modeling techniques the authors used as a form of power analysis, and I am unfamiliar with

Bayesian statistics. Also, I am unfamiliar with molecular genetics analyses. Finally, I have never conducted the sorts of behavioral assays that form the core of this research.

This is my single substantial concern:

Q3, P3 - "Most MaxEnt papers use cross-validation and the area under the curve (AUC) to evaluate model performance."

For the pre-registration to constrain researcher degrees of freedom, you need to state either (1) that you will use this method or (2) the decision rule you will use to determine whether you will use this method (and what you would do instead).

Author's reply:

Dear Dr.'s González, Nieberding, Parker, and an anonymous reviewer,

We are so glad that you thought we did a good job with the revision! Indeed, all of your comments so greatly improved our manuscript! We are happy to make the final revision and prepare it for in principle recommendation.

We revised our preregistration at <http://corinalogan.com/Preregistrations/gxpopbehaviorhabitat.html>, and we responded to your two 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/gxpopbehaviorhabitat.Rmd>. In case you want to see the history of track changes for this document at GitHub, click the link, then click the "History" button (right near 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 very much for your time!

Sincerely,

Corina, Kelsey, Alexis, Nancy, and Dieter

Reviewed by Caroline Marie Jeanne Yvonne Nieberding, 2020-09-18 07:44

Comment 1: Review 2 of revised version of manuscript by Logan et al "Implementing a rapid geographic range expansion - the role of behavior and habitat changes" submitted to PCI Ecology

I would like to congratulate the authors for revising so thoroughly their manuscript, and I have no further comments or concerns after reading the responses to referees and revised manuscript. Except for one useful "detail" regarding comment and response 39: in fact, the actual original paper showing that learning is costly was by Mery and Kawecki in Science in 2005. Using experimental evolution on learning for oviposition in flies, they showed that learning induced a reduction in offspring production and in survival, as far as I remember.

Best wishes for collecting sufficient data, CN.

Response 1: Thank you so much for your help and for your congratulations on our revision! Thank you also for highlighting the impact of the Mery and Kawecki (2005) article. We added new sentences to the end of the paragraph to incorporate this key example: Prediction 1 > Alternative 1: "Both of these alternatives assume that learning is costly [e.g., @mery2005cost], therefore individuals avoid it if they can. In the first case, individuals might not need to rely much on learning because they are attending to familiar cues across their range, therefore they only need to learn where in this new space these cues are located. In the second case, individual learning that the founding individuals needed to rely on to move into this new space could have been lost due to potential pressure to reduce this investment as soon as possible after moving to a new location."

Reviewed by anonymous reviewer, 2020-10-02 15:56

Comment 2: Thank you for inviting me to review this revised pre-print manuscript. I think the authors did a great job in the revision – the investigation goal presented in the Introduction is clearer than the previous version, the information presented in the current revision is also consistent with the goal of the investigation. I understand that there are many areas to be explored in the study, but the authors have addressed my previous comments and concerns appropriately. Hence, there is no major issue to be raised.

Response 2: We are so glad you think this version is clearer and more consistent. Thank you very much for your help!

Reviewed by Tim Parker, 2020-09-18 18:34

Comment 3: This pre-registration draft is a plan for studying range expansion in great-tailed grackles. The authors present clear questions and predictions, and detailed analysis plans. This is an appropriate pre-registration.

This authors of this pre-registration have addressed nearly all the concerns I laid out in my review of the prior draft. I have only one recommendation for a change prior to archiving (see below)

However, as I stated in the original review, I wish to acknowledge that I lack expertise regarding some of the methods in this pre-registration, and therefore cannot attest to their sufficiency. In particular, I am unfamiliar with the modeling techniques the authors used as a form of power analysis, and I am unfamiliar with Bayesian statistics. Also, I am unfamiliar with molecular genetics analyses. Finally, I have never conducted the sorts of behavioral assays that form the core of this research.

This is my single substantial concern:

Q3, P3 - “Most MaxEnt papers use cross-validation and the area under the curve (AUC) to evaluate model performance.”

For the pre-registration to constrain researcher degrees of freedom, you need to state either (1) that you will use this method or (2) the decision rule you will use to determine whether you will use this method (and what you would do instead).

Response 3: We are so glad to hear that we were able to address almost all of your concerns! Thank you for bringing up this further point, which allows us to better clarify our analysis plan. We made the following addition in bold:

Analysis Plan > Q3 Habitat > “Most MaxEnt papers use cross-validation and the area under the curve (AUC) to evaluate model performance, **and we will do the same.**”

Revision round #1

2020-08-11

I have now received the comments from 3 experienced reviewers on your preprint. The three of them think that your preprint is of interest and that you have made a great effort on putting it together, but they all include many comments that can help to improve it. Therefore, I am going to ask you to have a deep look to all the issues raised by the reviewers and submit a revised version of it.

Preprint DOI: <http://corinalogan.com/Preregistrations/gxpopbehaviorhabitat.html>

Reviewed by [Pizza Ka Yee Chow](#), 2020-07-14 07:27

I have reviewed Logan and colleagues' preregistered manuscript title 'Implementing a rapid geographic range expansion - the role of behavior and habitat changes'. The authors would like to examine the role of behaviours and habitat suitability in relation to an invasive species expansion, using Great-tailed grackles (*Quiscalus mexicanus*) as study species. To do so, they will assess multiple behaviours that have been shown or are thought to related to range expansion using several tasks (e.g. behavioural flexibility innovation, reversal learning, exploration) alongside dispersal behaviours within several populations at different stage of expansion. The authors will also include habitat-related variables (e.g. availability, suitability) in their investigation.

I think this work is important; not many studies to date have covered both internal factors such as characteristics (behaviour) of a species and external factors (habitat suitability/availability) in relation to species expansion. This study will help to shed lights on factors related to invasion success or successful settlement in new environments. While I find the study concept is important and worth to be investigated, I also find there are some major issues and queries in relation to smaller aspects within the concept (see below). Perhaps, this is down to the authors have provided a very brief version of their study (i.e. pre-registration). In this review, I have provided some suggestions here, which I hope they would be help the authors to refine their study design and write up for the final submission.

1) The abstract provides a very brief study background and the study objectives. However, it does not convey clearly the idea of the alternative explanation for range expansion. One issue here is that having suitable habitats as a facilitator of an species expansion is not new. In particular in ecology and more specifically invasive ecology, be it alone in plants or animals. Yet, there is no reference to support such idea.

2) Another major issue here is down to what the authors would like to do: are the authors seeking behaviour OR habitat suitability is the cause of range expansion? Or are they examining the relative roles of behaviour AND habitat suitability? (as the authors have stated in the C) Hypothesis : 'the relative roles of changes in behaviour and changes in habitats in the range expansion of great-tailed grackles.'). The former question appears to argue either nature or nurture whereas the latter is more prone to a combination of both. In the actual write up, the authors should clarify this concept succinctly.

3) Hypothesis: It is good that the authors are looking at several behaviours to understand the research question. However, the authors appear to weight up all behaviours in understanding the research question. Indeed, any two behaviours may vary their importance within the same expansion stage. For example, looking each trait at within population level, exploration may be more important than flexibility (despite both traits may be correlated in some ways) within the 'edge' populations (and not only between 'edge' and 'recently established' populations) because grackles may have to secure resources (e.g. places to stay, food to eat etc). That is to say each behaviour of interest may relate to the stage of range expansion differently and the authors should have different predictions for each behaviour. The lots-of-perhaps in the Hypothesis section may provide explanation for populations at different expansion stage, but the importance of behavioural traits may vary and shall be understood in populations that are at the same stage of expansion.

4) this comment is related to comment 2 - Assuming the authors are not testing 'either-or' but 'relatively importance'. When we talk about the relatively role, I think hypothesis should be stated in a way that should reflect the relative proportion of each role in the process. For example, H1 shall be 'if behaviour plays a more important role than habitat-related factors in expansion'(?).

Others

Abstract 1) Clarity –I suggest the authors write it clearly or provide more informative labels for each study population (e.g. the population in the centre is 'recently published' and the edge of the population is 'invasive front population'); the label will allow the readers to know it right away that the authors are comparing populations at the front of expansion with those that are established or at the middle of expansion.

2) Habitat availability? Or 'suitability'? the two words mean very different things and different measurements, please clarify.

Hypotheses 1) This hypothesis is needed to be more precise in that new location and range expansion could be seen in two ways but could well be depending on how authors are measuring these things. Are the authors stating the new locations where the grackles invade is a geographically continuous landscape? Or a completely different locations? 'Expansion' implies it is the former case. If this is true, the continuous landscape may not completely pose a higher challenge to grackles that lead higher behaviour flexibility than recently established population.

Protocols and open materials 1) Thank you for providing a detailed protocol – I have read through them point-by-point. The design of each task is either adopted from other studies or established set up for grackles. However, there are some key information missing here. For example, the authors may state clearly that all grackles will go through a habituation period novel apparatus (details of habituation period could be find online – what the authors have provided for this pre-registration); this will allow readers to know that task performance and participation rate presumably would be not be affected by neophobia but more down to motivation or other reasons (e.g. weather).

2) Provided information for each task is not entirely consistent throughout the section. For example, total duration of assessing exploration is provided but not in flexibility or innovativeness tasks.

3) The rationales of some measurements are not entirely clear to me. For example, why would the authors need to analyse DNA of grackles, how does the relatedness related to the research question?

4) Suitable habitat: please name a few ecological variables that are important for grackles. Are the authors using some kinds of index to indicate the degree of suitability of a habitat?

5) Flexibility task: how long does it session last for?

6) Innovativeness: What is the flexibility measure for this task? When a bird has successfully used a solution to solve the task, why did the authors block the previously successful solution and not allow the bird to explore an alternative solution? My two cents is there are pros and cons here, if the authors allow a bird to explore new solution, this would be a way to measure natural exploration tendency (which is another variable that the authors are interested in).

7) Exploration – the authors would like to go for simplicity by testing novel object and not novel environment, but what if exploration of novel object and environments correlate with boldness in opposite direction? Also, regarding relevance to what the authors are interested in, one shall assume exploring new environment test (which may allow invasive species to explore novel resources/ 'object') would be related to invasive species expansion. How would the authors measure exploration here? By the frequent of manipulating the object or the duration? I do not get this until I read the analysis plan.

8) Persistence. It is good that the authors have given habituation period for grackles as well as having a relatively strict passing criteria to ensure neophobia would not be a confound for task performance and participation. A note is that the authors may want to clarify why the proportion of trials participated in the flexibility and innovativeness reflect 'persistence' – I cannot get my head around this...as the measure could equally reflect 'high motivation' or 'eagerness to participate in task'.

E. Analysis Plan Model and simulation I agree with the authors that using hypothesis-appropriate mathematical model is a good way to analyse the data. A note on the analyses plan is that although the authors may set prior distribution from available, or the authors', publications, the authors may want to incorporate a larger and smaller mean and SD to increase the robustness of the results (i.e. to reflect whether the results in the current study is covered within the probability distribution).

Reviewed by [Tim Parker](#), 2020-07-30 19:27

This pre-registration draft is a plan for studying range expansion in great-tailed grackles. The authors present relatively clear hypotheses and predictions, and detailed analysis plans.

It is my opinion that, as a pre-registration, this draft is almost ready to be archived, although I have some specific suggestions for improvement. For the most part, the methods are presented clearly and with a high degree of detail (except for H3). Also, to the extent that my expertise allows me to evaluate the methods, those methods appear reasonable.

However, I wish to acknowledge that I lack expertise regarding some of the methods in this pre-registration, and therefore cannot attest to their sufficiency. In particular, I am unfamiliar with the modeling techniques the authors used as a form of power analysis, and I am unfamiliar with Bayesian statistics. Also, I am unfamiliar with molecular genetics analyses. Finally, I have never conducted the sorts of behavioral assays that form the core of this research.

Hypothesis – Predictions framework:

Because the authors have chosen to present a framework of hypotheses and predictions, I feel compelled point out that they have not used this framework in the traditional manner, and so I found their use of the framework confusing. This is a bit of a pet issue with me, so I apologize in advance for what follows, but I do very much believe that the tendency for the community of evolutionary biologists and ecologists to not rigorously follow the hypothesis-prediction framework when it is invoked hinders understanding and clarity of thinking.

Traditionally, a hypothesis is a tentative statement regarding how the world works, and a prediction of that hypothesis is something that the scientist should be able to observe if the hypothesis is true. Therefore, if the researcher examines the prediction and finds a lack of evidence for it, this should undermine confidence in the hypothesis. Thus a prediction is just a statement of what the researcher should observe/measure given a hypothesis is correct, and a hypothesis cannot have conflicting predictions. If you have constructed conflicting predictions, that is a sign that you have multiple (alternative) hypotheses.

For instance, the way Prediction 1 and Prediction 1 alternative 1 for H1 are presented confused me. I thought you were presenting two divergent (partly conflicting) predictions for the same hypothesis. However, after looking at Fig 1, I decided that 'Prediction 1 alternative' was maybe supposed to be a prediction of H3 (though this prediction as currently worded is not an ideal prediction of H3 as currently worded). Anyway, below is what I wrote in response to that paragraph before I looked at Fig 1. I'm including it here because I hope it will help you recognize my confusion and will help you clarify how you present hypothesis and predictions. I encourage you to re-work your descriptions of all your hypotheses and predictions so that they adhere to the standard framework.

Prediction 1 and Prediction 1 alternative 1 for H1 are in essence two different hypotheses (in part). One hypothesis is something like: the range expansion in great tailed grackles is facilitated by behavioral traits (flexibility, innovation, exploration, and persistence [actually, each of these should probably be considered a separate hypothesis]) that are found disproportionately at the leading edge of the range expansion. The other hypothesis is something like: the range expansion in great tailed grackles is facilitated by behavioral traits (flexibility, innovation, exploration, and persistence) that are characteristic of this species. You could divide up these hypotheses in other ways, but the point is that the predictions for the 1st half of both of these hypotheses are identical (presence of behavioral flexibility/ innovation/ exploration/ persistence at the leading edge), but the predictions for the 2nd parts of both of these hypotheses are different (behavioral flexibility/ innovation/ exploration/ persistence greater at leading edge vs. spread evenly through the entire population).

In a pre-registration, clarity about hypothesis and predictions is useful, because this allows the researchers to clearly state what they will conclude about each separate (component of their) hypothesis based on the outcome of each separate prediction.

protocols: Why not include the detailed protocols for H1 (now in a separate Google Doc) as part of the pre-registration?

Flexibility: Under what condition would you decide to “modify this protocol by moving the passing criterion sliding window in 1-trial increments, rather than 10-trial increments”?

Blinding during analyses: Would you like to present any justification for your lack of blinding?

Analysis Plan, H1

As I understand it, you present a clear criterion for statistical decisions (“From the pairwise contrasts, if the difference between the distributions crosses zero (yes), then we are not able to detect differences between the two sites. If they do not cross zero (no), then we are able to detect differences between the two sites.”)

However, a bit more explanation here for those not familiar with your analytical methods would be welcome.

Analysis Plan, H2

Is there only one value for ‘relatedness’ produced by this method? in other words, is their undisclosed analytical flexibility here?

Analysis Plan, H3

This appears to be the weakest part of the pre-registration (the vaguest portion, and thus the portion for which this pre-registration does not appear to be doing the work of constraining analytical options and thus constraining ‘researcher degrees of freedom’)

Can you provide more information about some of your explanatory variables? What exactly will the climate variables be? How will predator density be measured? Can you explain 'Distance to the next suitable habitat patch weighted by nearest mountain range/forest'? How will you define 'conspecific population' (for explanatory variable #6)? Will it be the detection of any individuals, or the detection of some minimum number of individuals? Can you provide any more info about your decision making process while fitting models using maxent?

Trivial comments:

Typo in abstract: "We first aim to compare behavior in wild-caught grackled"

Reviewed by [Caroline Marie Jeanne Yvonne Nieberding](#), 2020-08-11 10:45

Dear Authors, please find in attachment my comments to your proposed research project. Overall it is very interesting and well thought; some of my comments end up being due to finding the place where you produce the information I was looking for. Hopefully some comments will be useful to further improve the link between your experimental work and their relevance to the ecology of the species. Good luck with the covid crisis, Best regards, C. Nieberding.

[Download the review \(PDF file\)](#)

Author's reply:

Dear Dr.'s González, Chow, Parker, and Nieberding,

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 happy to have the opportunity to revise and resubmit.

We revised our preregistration at <http://corinalogan.com/Preregistrations/gxpopbehaviorhabitat.html>, and we responded to your comments below.

Note that the version-tracked version of this preregistration is in markdown at GitHub: <https://github.com/corinalogan/grackles/blob/master/Files/Preregistrations/gxpopbehaviorhabitat.Rmd>. In case you want to see the history of track changes for this document at GitHub, click the link, then click the "History" button (right near 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!

Two additional things: Due to COVID-19 issues, we have had to delay our data collection start date by a month. We now plan to begin collecting data in mid-October. We added a new co-author, Alexis Breen, who just joined the grackle team.

Thank you very much for your time!

Sincerely,

Corina, Kelsey, Alexis, Nancy, and Dieter

Round #1

Your decision

by Esther Sebastián González, 2020-08-11 13:40

Manuscript: <http://corinalogan.com/Preregistrations/gxpopbehaviorhabitat.html>

Please revise your pre-print

COMMENT 1: I have now received the comments from 3 experienced reviewers on your preprint. The three of them think that your preprint is of interest and that you have made a great effort on putting it together, but they all include many comments that can help to improve it. Therefore, I am going to ask you to have a deep look to all the issues raised by the reviewers and submit a revised version of it.

RESPONSE 1: Thank you very much for facilitating this process! We responded to all comments below.

Reviews Reviewed by Pizza Ka Yee Chow, 2020-07-14 07:27

I have reviewed Logan and colleagues' preregistered manuscript title 'Implementing a rapid geographic range expansion - the role of behavior and habitat changes'. The authors would like to examine the role of behaviours and habitat suitability in relation to an invasive species expansion, using Great-tailed grackles (*Quiscalus mexicanus*) as study species. To do so, they will assess multiple behaviours that have been shown or are thought to related to range expansion using several tasks (e.g. behavioural flexibility innovation, reversal learning, exploration) alongside dispersal behaviours within several populations at different stage of expansion. The authors will also include habitat-related variables (e.g. availability, suitability) in their investigation.

COMMENT 2: I think this work is important; not many studies to date have covered both internal factors such as characteristics (behaviour) of a species and external factors (habitat suitability/availability) in relation to species expansion. This study will help to shed lights on factors related to invasion success or successful settlement in new environments. While I find the study concept is important and worth to be investigated, I also find there are some major issues and queries in relation to smaller aspects within the concept (see below). Perhaps, this is down to the authors have provided a very brief version of their study (i.e. pre-registration). In this review, I have provided some suggestions here, which I hope they would be help the authors to refine their study design and write up for the final submission.

RESPONSE 2: Thank you very much for your positive feedback and for providing comments on how we can improve this work! We look forward to addressing your detailed comments below.

COMMENT 3: 1) The abstract provides a very brief study background and the study objectives. However, it does not convey clearly the idea of the alternative explanation for range expansion. One issue here is that having suitable habitats as a facilitator of an species expansion is not new. In particular in ecology and more specifically invasive ecology, be it alone in plants or animals. Yet, there is no reference to support such idea.

RESPONSE 3: Great point! We added citations to the Abstract and to the new Introduction for the alternative.

COMMENT 4: 2) Another major issue here is down to what the authors would like to do: are the authors seeking behaviour OR habitat suitability is the cause of range expansion? Or are they examining the relative roles of behaviour AND habitat suitability? (as the authors have stated in the C) Hypothesis : 'the relative roles of changes in behaviour and changes in habitats in the range expansion of great-tailed grackles.'). The former question appears to argue either nature or nurture whereas the latter is more prone to a combination of both. In the actual write up, the authors should clarify this concept succinctly.

RESPONSE 4: This is a really good point and one we need to clarify. Thanks to your comment, we realized that we were giving mixed messages, but in reality, we are not able to compare habitat and behavior with each other because the data for these variables are being collected at completely different scales and not on the same individuals. We are testing habit and behavior individually to assess whether either or both play a role in the range expansion. We updated Figure 1 to clarify this point, and we modified Figure 3 and added Figure 4 to help show this. We also clarified this in the text in the following sections:

Abstract: “However, it is an alternative non-exclusive possibility that an increase in the amount of available habitat can be a facilitator for a range expansion.”

The last sentence of the Abstract and Introduction: “Results will elucidate whether the rapid geographic range expansion of great-tailed grackles is associated with individuals differentially expressing particular behaviors **and/or** whether the expansion is facilitated by the alignment of their natural behaviors with an increase in suitable habitat (i.e., human-modified environments).”

Hypotheses (the note at the top, which is now in the Introduction): “There could be multiple mechanisms underpinning the results we find, however our aim here is to narrow down the **role** of changes in behavior and changes in habitats in the range expansion of great-tailed grackles”

COMMENT 5: 3) Hypothesis: It is good that the authors are looking at several behaviours to understand the research question. However, the authors appear to weight up all behaviours in understanding the research question. Indeed, any two behaviours may vary their importance within the same expansion stage. For example, looking each trait at within population level, exploration may be more important than flexibility (despite both traits may be correlated in some ways) within the ‘edge’ populations (and not only between ‘edge’ and ‘recently established’ populations) because grackles may have to secure resources (e.g. places to stay, food to eat etc). That is to say each behaviour of interest may relate to the stage of range expansion differently and the authors should have different predictions for each behaviour. The lots-of-perhaps in the Hypothesis section may provide explanation for populations at different expansion stage, but the importance of behavioural traits may vary and shall be understood in populations that are at the same stage of expansion.

RESPONSE 5: Thank you for pointing out that it was unclear that we are only investigating whether these behaviors are important at the very edge. We have now removed the note from the Hypothesis section, and expanded on it in the new Introduction, which we hope will make the “perhaps” in the predictions make more sense.

Introduction: “It is generally thought that behavioral flexibility, the ability to change behavior when circumstances change (see @mikhalevichis2017 for theoretical background on our flexibility definition), plays an important role in the ability of a species to rapidly expand their geographic range (e.g., @lefebvre1997feeding, @griffin2014innovation, @chow2016practice, @sol2000behavioural, @sol2002behavioural, @sol2005big, @sol2007big). These ideas predict that flexibility, exploration, and innovation facilitate the expansion of individuals into completely new areas and that their role diminishes after a certain number of generations [@wright2010behavioral]. In support of this, experimental studies have shown that latent abilities are primarily expressed in a time of need [e.g., @taylor2007spontaneous; @bird2009insightful; @manrique2011spontaneous; @auersperg2012spontaneous; @laumer2018spontaneous]. Therefore, we do not expect the founding individuals who initially dispersed out of their original range to have unique behavioral characteristics that are passed on to their offspring. Instead, we expect that the actual act of continuing a range expansion relies on flexibility, exploration, innovation, and persistence, and that these behaviors are therefore expressed more on the edge of the expansion range where there have not been many generations to accumulate relevant knowledge about the environment.”

COMMENT 6: 4) this comment is related to comment 2 - Assuming the authors are not testing 'either-or' but 'relatively importance'. When we talk about the relatively role, I think hypothesis should be stated in a way that should reflect the relative proportion of each role in the process. For example, H1 shall be 'if behaviour plays a more important role than habitat-related factors in expansion'(?).

RESPONSE 6: We hope that our revision in Response 4 clarified that we are not examining the relative roles of habitat and behavior with each other, but each one separately.

Others

COMMENT 7: Abstract 1) Clarity –I suggest the authors write it clearly or provide more informative labels for each study population (e.g. the population in the centre is ‘recently published’ and the edge of the

population is ‘invasive front population’); the label will allow the readers to know it right away that the authors are comparing populations at the front of expansion with those that are established or at the middle of expansion.

RESPONSE 7: Thank you for pointing this out. We avoid using the word “invasion” with this species because invasion ecologists think that a species must be introduced by humans for it to be invasive. Great-tailed grackles have primarily introduced themselves across their range, therefore we tend to avoid the invasion term. We added study location points to the map in Figure 3 and we revised the text as follows to improve clarity: Abstract: “core of the original range, a more recent population in the middle of the northern expansion front, a very recent population on the northern edge of the expansion front”

COMMENT 8: 2) Habitat availability? Or ‘suitability’? the two words mean very different things and different measurements, please clarify.

RESPONSE 8: Good point that we were unclear about these terms. We now defined them in the Abstract and in the new Introduction.

Abstract: “However, it is an alternative non-exclusive possibility that an increase in the amount of available habitat can be a facilitator for a range expansion.”

Abstract: “3) these species use different habitats, habitat suitability and connectivity (which combined determines whether habitat is available) has increased across their range, and what proportion of suitable habitat both species occupy.”

COMMENT 9: Hypotheses 1) This hypothesis is needed to be more precise in that new location and range expansion could be seen in two ways but could well be depending on how authors are measuring these things. Are the authors stating the new locations where the grackles invade is a geographically continuous landscape? Or a completely different locations? ‘Expansion’ implies it is the former case. If this is true, the continuous landscape may not completely pose a higher challenge to grackles that lead higher behaviour flexibility than recently established population.

RESPONSE 9: The expansion is occurring in a geographically continuous landscape, therefore we are looking at a relative difference between populations. We consider the Woodland population to be close enough to the range edge to be classified as an edge population because there have been so few generations there that individuals are likely to still encounter new elements in their environment that they could not have learned about socially. Woodland, California is not the northernmost part of the great-tailed grackle’s range, however it is as far north as we could go while operating under the constraint that we need a large enough population that exists there year round to be able to conduct such a study as this. We clarified as follows:

Introduction: “Instead, we are investigating whether the actual act of continuing a range expansion relies on flexibility, exploration, innovation, and persistence, and that these behaviors are therefore expressed more on the edge of the expansion range where there have not been many generations to accumulate relevant knowledge about the environment.”

COMMENT 10: Protocols and open materials 1) Thank you for providing a detailed protocol – I have read through them point-by-point. The design of each task is either adopted from other studies or established set up for grackles. However, there are some key information missing here. For example, the authors may state clearly that all grackles will go through a habituation period novel apparatus (details of habituation period could be find online – what the authors have provided for this pre-registration); this will allow readers to know that task performance and participation rate presumably would be not be affected by neophobia but more down to motivation or other reasons (e.g. weather).

RESPONSE 10: Good point, thank you! We added to this section that we conduct habituation first for both the flexibility and innovativeness experiments and we added your point to the persistence description. We revised as follows:

Protocols and open materials > Flexibility: “Grackles are first habituated to a yellow tube and trained to search for hidden food.”

Protocols and open materials > Flexibility: “Grackles are first habituated to the log apparatus with all of the doors locked open and food inside each locus. After habituation, the log, which has four ways of accessing food...”

Protocols and open materials > Persistence: “Persistence is measured as the proportion of trials participated in during the flexibility and innovativeness experiments (after habituation, thus it is not confounded with neophobia)”

COMMENT 11: 2) Provided information for each task is not entirely consistent throughout the section. For example, total duration of assessing exploration is provided but not in flexibility or innovativeness tasks.

RESPONSE 11: This is because only the exploration assay has a set session duration because the aim is to determine the latency to approach a novel object. To make sure results are comparable across birds, the session has to be standardized across birds so they all get the same amount of time with the apparatus. For the reversal learning and multiaccess log experiments, we only pay attention to when they pass criterion, therefore sessions last for varying amounts of time depending on the grackle’s motivation. In the protocol, we provide only the information that the experimenter needs to attend to when testing the birds, so what is relevant for one task might not be relevant or noteworthy for another task.

COMMENT 12: 3) The rationales of some measurements are not entirely clear to me. For example, why would the authors need to analyse DNA of grackles, how does the relatedness related to the research question?

RESPONSE 12: This is a great point, sorry for the confusion! We now clarify this in the new Introduction.

COMMENT 13: 4) Suitable habitat: please name a few ecological variables that are important for grackles. Are the authors using some kinds of index to indicate the degree of suitability of a habitat?

RESPONSE 13: The model we will run on this data, MaxEnt, produces a continuous prediction of habitat suitability for each grid cell (0 is least suitable and 1 is most suitable). We will also use jackknifing procedures to evaluate the relative contribution/importance of different environmental variables to the probability of species occurrence. We added this clarification to Analysis Plan > Q3 > P3 > Explanatory variables.

We added detailed descriptions of each variable in the Analysis Plan > Q3 (please see Response 29 for the changes), and we provided examples of these variables in the revised text in the new Introduction and also in:

Protocols and open materials > Suitable habitat: “We identified suitable habitat variables from Selander and Giller (1961), Johnson and Peer (2001), and Post et al. (1996) (e.g., types of suitable land cover including wetlands, marine coastal, arable land, grassland, mangrove, urban), and we added additional variables relevant to our hypotheses (e.g., distance to nearest uninhabited suitable habitat patch to the north, presence/absence of water in the area).”

COMMENT 14: 5) Flexibility task: how long does it session last for?

RESPONSE 14: Please see Response 11.

COMMENT 15: 6) Innovativeness: What is the flexibility measure for this task? When a bird has successfully used a solution to solve the task, why did the authors block the previously successful solution and not allow the bird to explore an alternative solution? My two cents is there are pros and cons here, if the authors allow a bird to explore new solution, this would be a way to measure natural exploration tendency (which is another variable that the authors are interested in).

RESPONSE 15: For the multiaccess box, we follow the methods developed by Auersperg et al. (2011). We and Auersperg et al. (2011) interpreted the switching between options as a measure of flexibility instead of

exploration. This fits with our definition of flexibility, which is to decide among a variety of options which choice is a functional option to attempt. We can see how exploration tendency might come into play the first time a bird has touched an option, however the box is not new to the bird by the time the test starts because they undergo habituation with it with the doors locked open. Your idea to measure exploration on the log would work well if one were to video record the habituation period and measure latency to first touch to each locus. In this case, it would be a great measure of exploration if the presence of food in all loci was acceptable.

Although in a previous study, we extracted a flexibility measure from the multiaccess box task (the latency to attempt to solve a new locus after having previously successfully solved a different locus; Logan et al. 2019 http://corinalogan.com/Preregistrations/g_flexmanip.html), we are not going to examine flexibility in the context of the multiaccess box task in the current study. For flexibility, we will use one measure: the number of trials to reverse a color preference. Sorry if this was confusing. Please let us know if there is a place in the text that we need to clarify.

The reason for blocking off an option that a bird previously demonstrated proficiency on is to force them to try to solve the other options because we are interested in how many options a bird can solve. If we didn't block off a previously successful option, then the bird might only use that one option to repeatedly obtain the food and we would not have an accurate measure of their innovative potential. There was only one way to solve each option, so once they were proficient at a particular locus, there wouldn't have been any alternative ways of solving that locus. We now mention that our experimental design is after Auersperg in Protocols and open materials > Innovativeness.

COMMENT 16: 7) Exploration – the authors would like to go for simplicity by testing novel object and not novel environment, but what if exploration of novel object and environments correlate with boldness in opposite direction? Also, regarding relevance to what the authors are interested in, one shall assume exploring new environment test (which may allow invasive species to explore novel resources/ 'object') would be related to invasive species expansion. How would the authors measure exploration here? By the frequent of manipulating the object or the duration? I do not get this until I read the analysis plan.

RESPONSE 16: This is a great question and one that we will have answers to soon! We are currently analyzing data from a study we conducted in Arizona grackles that looks at the relationship between novel object and novel environment performance as well as whether performance on both relates with boldness all in the same individuals (McCune et al. 2019 http://corinalogan.com/Preregistrations/g_exploration.html). Once we know these relationships, we will be able to decide which exploration test (or tests) to include in the current study and explain if and how they relate to boldness.

In terms of whether exploration of a novel environment/object relates to the exploration of novel resources/objects in the wild, we investigate these links in a separate preregistration: space use (<http://corinalogan.com/Preregistrations/gspaceuse.html>). In space use, our analysis for H1 (across all three populations) will tell us whether exploration of novel object/environment is correlated with space use in the wild. If these variables are correlated we will be able to infer that results from the analyses of movement behavior across populations (space use H2) likely also apply to the exploration of the novel object/environment in the aviary assays.

Sorry that the methods for exploration weren't clear until the end! We placed a detailed explanation of the methods in Methods > Protocols and open materials, and we summarized the method in Prediction 1, as well as describing the method in the new Introduction. We hope that this helps readers understand how the test works sooner in the text.

COMMENT 17: 8) Persistence. It is good that the authors have given habituation period for grackles as well as having a relatively strict passing criteria to ensure neophobia would not be a confound for task performance and participation. A note is that the authors may want to clarify why the proportion of trials participated in the flexibility and innovativeness reflect 'persistence' – I cannot get my head around this...as the measure could equally reflect 'high motivation' or 'eagerness to participate in task'.

RESPONSE 17: We're glad you like the habituation passing criteria! We find that it works really well in practice to make sure we aren't testing neophobic birds. For the persistence measure, if a bird participated in 10/10 trials, then they would score a 1, and this would indicate they have a high participation level. Alternatively, if a bird participated in only 1/10 trials, then they would score a 0.1, and this would indicate that they did not persist in attempting to participate in trials. The lack of participation could be due in part to motivation, however we think motivation is impossible to measure in this species because limited food restriction often does not get them to participate in a trial if they really don't want to participate. We think that the birds who choose to participate more often could be more eager to participate in the task and in this sense I think we mean the same thing by eagerness and persistence. What we have noticed when testing grackles in Santa Barbara and Arizona is that, when all birds have equal opportunities to participate in trials every day, those birds who do not participate as often could be considered less persistent in terms of their persistence with engaging with the task. We added a clarification to:

Methods > Protocols and open materials > Persistence: "This measure indicates that those birds who do not participate as often are less persistent in terms of their persistence with engaging with the task."

COMMENT 18: E. Analysis Plan Model and simulation I agree with the authors that using hypothesis-appropriate mathematical model is a good way to analyse the data. A note on the analyses plan is that although the authors may set prior distribution from available, or the authors', publications, the authors may want to incorporate a larger and smaller mean and SD to increase the robustness of the results (i.e. to reflect whether the results in the current study is covered within the probability distribution).

RESPONSE 18: Yes, the information from a subset of one population might not reflect the variation found in the data we are going to collect. Therefore, we previously assessed whether the prior distribution we chose for the Bayesian analyses would cover a range of expected results in the study through prior simulations. We realize that we had included the code for the prior simulations, but not mentioned this in the text - sorry for the confusion! We now added the following:

Analysis Plan > Hypothesis-specific mathematical model: "We formulated these models in a Bayesian framework. We determined the priors for each model by performing prior predictive simulations based on ranges of values from the literature to check that the models are covering the likely range of results."

Reviewed by Tim Parker, 2020-07-30 19:27 This pre-registration draft is a plan for studying range expansion in great-tailed grackles. The authors present relatively clear hypotheses and predictions, and detailed analysis plans.

COMMENT 19: It is my opinion that, as a pre-registration, this draft is almost ready to be archived, although I have some specific suggestions for improvement. For the most part, the methods are presented clearly and with a high degree of detail (except for H3). Also, to the extent that my expertise allows me to evaluate the methods, those methods appear reasonable. However, I wish to acknowledge that I lack expertise regarding some of the methods in this pre-registration, and therefore cannot attest to their sufficiency. In particular, I am unfamiliar with the modeling techniques the authors used as a form of power analysis, and I am unfamiliar with Bayesian statistics. Also, I am unfamiliar with molecular genetics analyses. Finally, I have never conducted the sorts of behavioral assays that form the core of this research.

RESPONSE 19: Thank you very much for your feedback! We look forward to addressing your comments below.

COMMENT 20: Hypothesis – Predictions framework: Because the authors have chosen to present a framework of hypotheses and predictions, I feel compelled point out that they have not used this framework in the traditional manner, and so I found their use of the framework confusing. This is a bit of a pet issue with me, so I apologize in advance for what follows, but I do very much believe that the tendency for the community of evolutionary biologists and ecologists to not rigorously follow the hypothesis-prediction framework when it is invoked hinders understanding and clarity of thinking. Traditionally, a hypothesis is a tentative statement regarding how the world works, and a prediction of that hypothesis is something that the scientist should be able to observe if the hypothesis is true. Therefore, if the researcher examines the

prediction and finds a lack of evidence for it, this should undermine confidence in the hypothesis. Thus a prediction is just a statement of what the researcher should observe/measure given a hypothesis is correct, and a hypothesis cannot have conflicting predictions. If you have constructed conflicting predictions, that is a sign that you have multiple (alternative) hypotheses. For instance, the way Prediction 1 and Prediction 1 alternative 1 for H1 are presented confused me. I thought you were presenting two divergent (partly conflicting) predictions for the same hypothesis. However, after looking at Fig 1, I decided that 'Prediction 1 alternative' was maybe supposed to be a prediction of H3 (though this prediction as currently worded is not an ideal prediction of H3 as currently worded). Anyway, below is what I wrote in response to that paragraph before I looked at Fig 1. I'm including it here because I hope it will help you recognize my confusion and will help you clarify how you present hypothesis and predictions. I encourage you to re-work your descriptions of all your hypotheses and predictions so that they adhere to the standard framework. Prediction 1 and Prediction 1 alternative 1 for H1 are in essence two different hypotheses (in part). One hypothesis is something like: the range expansion in great tailed grackles is facilitated by behavioral traits (flexibility, innovation, exploration, and persistence [actually, each of these should probably be considered a separate hypothesis]) that are found disproportionately at the leading edge of the range expansion. The other hypothesis is something like: the range expansion in great tailed grackles is facilitated by behavioral traits (flexibility, innovation, exploration, and persistence) that are characteristic of this species. You could divide up these hypotheses in other ways, but the point is that the predictions for the 1st half of both of these hypotheses are identical (presence of behavioral flexibility/ innovation/ exploration/ persistence at the leading edge), but the predictions for the 2nd parts of both of these hypotheses are different (behavioral flexibility/ innovation/ exploration/ persistence greater at leading edge vs. spread evenly through the entire population). In a pre-registration, clarity about hypothesis and predictions is useful, because this allows the researchers to clearly state what they will conclude about each separate (component of their) hypothesis based on the outcome of each separate prediction.

RESPONSE 20: Thank you for bringing this up. We can see now how there was confusion in how we presented the hypotheses and predictions. To address this, we changed the Hypotheses to Research Questions, which allowed us to keep the outcomes neutral (e.g., the hypothesis is not that they facilitate the range expansion, which implies a positive outcome, but rather that the question is about whether there are behavior changes across populations). We then added for each Prediction what hypothesis would be supported. We also clearly marked the actual predictions (in bold) and what hypothesis would be supported (in italics) so it is more organized and helpful for readers to follow.

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 <https://ecology.peercommunityin.org/public/rec?id=65&reviews=True>): "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."

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.

COMMENT 21: protocols: Why not include the detailed protocols for H1 (now in a separate Google Doc) as part of the pre-registration?

RESPONSE 21: We like to include the protocols as a link to the google doc because this is the document that the experimenters use when testing. The experimenters update this document as exceptions and notes occur. If we keep the link to the version-tracked google doc, then everyone can see where we are at in the process and what has happened so far, rather than making it a static part of the preregistration. We consider the document at the link as part of the preregistration, however, if you prefer, we could copy and paste the H1 protocols as they are currently into the Methods section of the preregistration.

COMMENT 22: Flexibility: Under what condition would you decide to “modify this protocol by moving the passing criterion sliding window in 1-trial increments, rather than 10-trial increments”?

RESPONSE 22: This is a really good point and we can decide right now. We will go with the 1-trial increments because this makes more logical sense than analyzing in a sliding 10-trial-block window, which is a socially inherited tradition in the field of comparative cognition. We updated as follows:

Methods > Protocols and open materials > Flexibility: “An individual is considered to have a preference if it chose the rewarded option at least 85% of the time (17/20 correct) in the most recent 20 trials (with a minimum of 8 or 9 correct choices out of 10 on the two most recent sets of 10 trials). We use a sliding window in 1-trial increments to calculate whether they passed after their first 20 trials.”

COMMENT 23: Blinding during analyses: Would you like to present any justification for you lack of blinding?

RESPONSE 23: Thanks for your comment, which also made us remember that we actually do conduct some analyses with blind coders. We updated this section to:

Methods > Blinding during analyses: “Blinding is usually not involved in the final analyses because the experimenters collect the data (and therefore have seen some form of it) and run the analyses. Hypothesis- and data-blind video coders are recruited to conduct interobserver reliability of 20% of the videos for each experiment.”

We also included a new section that describes our interobserver reliability analyses in Analysis Plan > Interobserver reliability of dependent variables

COMMENT 24: Analysis Plan, H1: As I understand it, you present a clear criterion for statistical decisions (“From the pairwise contrasts, if the difference between the distributions crosses zero (yes), then we are not able to detect differences between the two sites. If they do not cross zero (no), then we are able to detect differences between the two sites.”) However, a bit more explanation here for those not familiar with your analytical methods would be welcome.

RESPONSE 24: We can see where more information, particularly in a step by step way, would be useful - thank you for pointing this out. We added more information to explain the approach as follows:

Analysis Plan > Q1 > Hypothesis specific mathematical model: “We will then perform pairwise contrasts to determine at what point we will be able to detect differences between sites by manipulating sample size, and $\$alpha\$$ means and standard deviations. Before running the simulations, we decided that a model would detect an effect if 89% of the difference between two sites is on the same side of zero (following @statrethinkingbook). **We are using a Bayesian approach, therefore comparisons are based on samples from the posterior distribution. We will draw 10,000 samples from the posterior distribution, where each sample will have an estimated mean for each population. For the first contrast, within each sample, we subtract the estimated mean of the edge population from the estimated mean of the core population. For the second contrast, we subtract the estimated mean of the edge population from the estimated mean of the middle population. For the third contrast, we subtract the estimated mean of the middle population**

from the estimated mean of the core population. We will now have samples of differences between all of the pairs of sites, which we can use to assess whether any site is systematically larger or smaller than the others. We will determine whether this is the case by estimating what percentage of each sample of differences is either larger or smaller than zero. For the first contrast, if 89% of the differences are larger than zero, then the core population has a larger mean. If 89% of the differences are smaller than zero, then the edge population has a larger mean.”

Analysis Plan > Q1 > Table 2 > legend: “Simulation outputs from varying sample size (n), and α means and standard deviations. We calculate pairwise contrasts between the estimated means from the posterior distribution: if for a large sample the difference is both positive and negative and crosses zero (yes), then we are not able to detect differences between the two sites. **If the differences between the means are all on one side of zero for 89% of the posterior samples (no), then we are able to detect differences between the two sites. We chose the 89% interval based on [statrethinkingbook].**”

COMMENT 25: Analysis Plan, H2: Is there only one value for ‘relatedness’ produced by this method? in other words, is their undisclosed analytical flexibility here?

RESPONSE 25: There are multiple ways to calculate relatedness among pairs of individuals from genotypic data. We were originally thinking that we would compare the validity and robustness of different ways of calculating relatedness based on our data, but we had not mentioned this in the preregistration. Since submitting this preregistration, we have now checked various estimators as part of a separate preregistration (Sevchik et al. 2019; http://corinalogan.com/Preregistrations/gdispersal_manuscript.html) on a subset of the Arizona data, which suggested that the estimator by Queller & Goodnight appears most appropriate for our data. As such, we will now use only the Queller & Goodnight method in the current preregistration. We clarified this as follows:

Analysis Plan > Q2 Dispersal: “Genetic relatedness between all pairs of individuals is calculated using the package “related” (@pew2015related) in R (as in @thrasher2018double) using the estimator by Queller & Goodnight, which was more robust for our inferences in a subset of the Arizona data [sevchik2019dispersal].”

COMMENT 26: Analysis Plan, H3: This appears to be the weakest part of the pre-registration (the vaguest portion, and thus the portion for which this pre-registration does not appear to be doing the work of constraining analytical options and thus constraining ‘researcher degrees of freedom’) Can you provide more information about some of your explanatory variables? What exactly will the climate variables be? How will predator density be measured? Can you explain ‘Distance to the next suitable habitat patch weighted by nearest mountain range/forest’? How will you define ‘conspecific population’ (for explanatory variable #6)? Will it be the detection of any individuals, or the detection of some minimum number of individuals? Can you provide any more info about your decision making process while fitting models using maxent?

RESPONSE 26: We now describe the reason for including each explanatory variable and what it might mean for a grackle range expansion, including explaining Distance to the next suitable habitat patch weighted by nearest mountain range/forest, and what the climate variables are (please see Response 29 for details). Our aim is not to precisely identify which variables are the primary constraints on where grackles can be found. Instead, we only want to identify suitable habitat across the Americas. Therefore, there is no decision making process in the model about which variables to include or not. We will optimize the model by trying different regularization coefficient values, which controls how much additional terms are penalized (Maxent’s way of protecting against overfitting), and choosing the value that maximizes model fit. Most MaxEnt papers use cross-validation and the area under the curve (AUC) to evaluate model performance. We added this description to Analysis Plan > Q3 > Explanatory variables.

COMMENT 27: Trivial comments: Typo in abstract: “We first aim to compare behavior in wild-caught grackled”

RESPONSE 27: Thank you for catching that! We fixed it!

Reviewed by Caroline Marie Jeanne Yvonne Nieberding, 2020-08-11 10:45

COMMENT 28: Dear Authors, please find in attachment my comments to your proposed research project. Overall it is very interesting and well thought; some of my comments end up being due to finding the place where you produce the information I was looking for. Hopefully some comments will be useful to further improve the link between your experimental work and their relevance to the ecology of the species. Good luck with the covid crisis, Best regards, C. Nieberding. Download the review (PDF file)

RESPONSE 28: We are so glad that you like the project! Thanks so much for your feedback (and also for the luck wishes during a time of COVID), which we include and respond to below. **COMMENT 29 (CN1):** Abstract: “3) these species use different habitats, habitat availability and connectivity” What habitat variables? This is tricky because: - the needs of the species need to be known (food sources, habitat type(s) for shelter and nest,...) - where do the data come from? this type of habitat information is not available through GIS / satellite / remote sensing? At the very end of the file I have found the list of specific habitat variables that you intend to map and compare to the occurrence data of the birds, but it would be a useful improvement to specify why/how these variables are important to the ecology of the species. So far it seems that you collect these habitat variables because they are available and they are not necessarily relevant to explain the species distribution. This is my major concern.

RESPONSE 29: Good points, thank you for pointing in the right direction in terms of what we needed to clarify. We now include an Introduction where we clarify for both species: 1) habitat types for foraging and nesting, 2) food sources, and 3) list examples of suitable habitat variables as well as describe variables that we added because they are hypothesis-relevant. Note that we removed predator density as an independent variable from the model because adult grackles have very few predators (i.e., two raptor species, one owl species, one snake species, and domestic cats; @johnson2001great) and predation is not noted in the literature as a major cause of mortality.

We also moved “Distance to the next suitable habitat patch weighted by nearest mountain range/forest” into the new “Distance between points on the northern edge of the range to the nearest uninhabited suitable habitat patch to the north in 1970 compared with the same patches in ~2018”, which replaced “Distance to the nearest conspecific population 10 years previous to the point in time being investigated”. These changes were made because we got clearer about what exactly the model is doing and what exactly we need to answer our questions. Thank you very much for your great questions which helped us narrow this down!

We also realized that we needed to pull out of the Land Cover variable the “distance from road/water body/wetland/water treatment plant” and move it into its own independent variable because it involves a separate treatment to obtain this data. It is now its own variable under “Presence/absence of water in the cell for each point”.

In the Analysis Plan we now describe the background for all variables (why we included them and what they could mean to a grackle range expansion) as follows:

Analysis Plan > Q3 > P3 > Explanatory Variables: “1) **Land cover** (e.g., forest, urban, arable land, pastureland, wetlands, marine coastal, grassland, mangrove) - we chose these land cover types because they represent the habitat types in which both species exist, as well as habitat types (e.g., forest) they are not expected to exist in [selander1961analysis] to confirm that this is the case. If it is the case, it is possible that large forested areas are barriers for the range expansion of one or both species. **We will download global land cover type data from MODIS (16 terrestrial habitat types) and/or the IUCN habitat classification (47 terrestrial habitat types).** The IUCN has assigned habitat classifications to great-tailed (<https://www.iucnredlist.org/species/22724308/132174807#habitat-ecology>) and boat-tailed (<https://www.iucnredlist.org/species/22724311/94859792#habitat-ecology>) grackle, however these appear to be out of date and we will update them for the purposes of this project.

2) **Elevation** - selander1961analysis notes the elevation range for GTGR (0-2134m), but not BTGR, therefore establishing the current elevation ranges for both species will allow us to determine whether and which mountain ranges present range expansion challenges. **We will obtain elevation data from USGS.**

3) **Climate** (e.g., daily/annual temperature range) - because this species was originally tropical [[@wehtje2003range](#)], which generally has a narrow daily and annual climate range, and now they exist in temperate regions, which have much larger climate ranges, this variable will allow us to determine potential climatic limits for both species. If there are limits, this could inform the difference between the range expansion rates of the two species. **We will consider the 19 bioclimatic variables from [WorldClim](#).**

4) **Presence/absence of water in the cell for each point** - both species are considered to be highly associated with water [e.g., [@selander1961analysis](#)], therefore we will identify how far from water each species can exist to determine whether it is a limiting factor in the range expansion of one or both species. The data will come from [USGS National Hydrography](#).

5) **Connectivity**: Distance between points on the northern edge of the range to the nearest uninhabited suitable habitat patch to the north in 1970 compared with the same patches in ~2018. We identified the northern edge of the distribution based on reports on eBird.org from 1968-1970, which resulted in recordings of GTGR in 48 patches and recordings of BTGR in 30 patches. For these patches, we calculated the connectivity (the least cost path) to the nearest uninhabited suitable habitat patch in 1970 and again in ~2018. Given that GTGR are not found in forests and that the elevation limits for GTGR [[@selander1961analysis](#)], and observing the sightings of both species on eBird.org, large forests, tall mountain ranges and high elevation geographic features could block or slow the expansion of one or both species into these areas and their surroundings. For each point, we will calculate the least cost path between it and the nearest location with grackle presence using the `leastcostpath` R package ([@leastcostpath](#)). This will allow us to determine the costs involved in a grackle deciding whether to fly around or over a mountain range/forest. We will define the forest and mountain ranges from the land cover and/or elevation maps.

COMMENT 30 (CN2): State of the data: “This preregistration was written (Mar 2020) prior to collecting any data from the edge and core populations.” Which means ? Please specify

RESPONSE 30: Good point. We clarified as follows: “This preregistration was written (Mar 2020) prior to collecting any data from the edge and core populations, therefore we were blind to these data”

COMMENT 31 (CN3): State of the data: “Some of the relatedness data from the middle population (Arizona) has already been analyzed for other purposes (n=57 individuals, see Sevchik et al. (2019)), therefore it will be considered secondary data: data that are in the process of being collected for other investigations. However, we have now collected blood samples from many more grackles in Arizona, therefore we will redo the analyses from the Arizona population in the analyses involved in the current preregistration” Relevant for this study ? Are you going to use relatedness/ genetic data ? Relevance unclear based on abstract above. It is clear later in the description of the protocols but it may be useful to be more explicit earlier about the type of data (snp) and that these data have been proven useful to quantify a range of relatedness in different populations of this species?

RESPONSE 31: Sorry for our lack of clarity! We now describe this in the new Introduction and we clarified the State of the Data as follows:

“However, we were not blind to some of the data from the Arizona population: some of the relatedness data (SNPs used for Hypothesis 2 to quantify relatedness to infer whether individuals disperse away from relatives) from the middle population (Arizona) has already been analyzed for other purposes (n=57 individuals, [see @sevchik2019dispersal](#)). Therefore, it will be considered secondary data: data that are in the process of being collected for other investigations. We have now collected blood samples from many more grackles in Arizona, therefore we will redo the analyses from the Arizona population in the analyses involved in the current preregistration.”

COMMENT 32 (CN4): State of the data: “This preregistration was submitted in May 2020 to PCI Ecology for pre-study peer review” I have not seen this data

RESPONSE 32: Sorry for the confusion. This just documents the time we submitted this preregistration to PCI Ecology, which is the submission you commented on. This section is more of a place for people to see which

parts of the study happened before data collection, after data collection and before data analysis, and after data analysis so readers can judge for themselves our level of bias throughout the process.

COMMENT 33 (CN5): State of the data: “**Level of data blindness:** Logan and McCune collect the behavioral data (H1) and therefore have seen this data for the Arizona population. Lukas has access to the Arizona data and has seen some of the summaries in presentations. Chen has not seen any data.” I think that this is not what is expected as an answer : behavioural studies may be biased mostly by knowing which outcome is expected for the animal at the time the data is collected. So we rather expect that the scientist who collected the behavioural observations is not aware of the population origin of the animals, and that animals from different populations (as far as possible) are randomized during successive observations. I understand that this may not be feasible for a study of such large geographical scale (as you write late in the protocols).

RESPONSE 33: Yes, in our case, the behavioral data are collected at the location the particular population is at, so experimenters always know which population they are working in. In registered reports, the level of data blindness is important to document because seeing the data can influence their future predictions about a particular question (see Chambers & Tzavella 2020 <https://osf.io/preprints/metaarxiv/43298/> for more details). We wanted to be clear up front what our potential for bias is.

COMMENT 34 (CN6): H1 > P1 “speed at reversing a previously learned color preference” For food items ?

RESPONSE 34: Yes, exactly. We clarified, thank you! We revised it to:

“speed at reversing a previously learned color preference based on it being associated with a food reward”

COMMENT 35 (CN7): H1 > P1 “innovativeness: number of options solved on a puzzle box” Link to natural selection in the wild is unclear ?

RESPONSE 35: We have now clarified this in the new Introduction.

COMMENT 36 (CN8): H1 > P1 “Perhaps in newly established populations, individuals need to learn about and innovate new foraging techniques or find new food sources” Relevant but then better link your experimental tests to ecologically relevant hypotheses. For example why test for learning of colour change, if not for food? Or puzzle tests while localization /exploration of new / scattered food item is perhaps more relevant? In general, this becomes more clear after one has read the protocols below, but this is my second and last real concern about this project: can you link better the expected ecological needs (for finding food) and the type of tests that you conduct here? To give you an example, in butterflies we test the specific host plant that females use to oviposit, and the test is about the time they need to find, and remember, the location of the host plant. The link to the demography and on selection in the field is more immediate.

RESPONSE 36: We have now clarified this in the new Introduction.

COMMENT 37 (CN9): H1 > P1 “Higher variances in behavioral traits indicate that there is a larger diversity of individuals in the population, which means that there is a higher chance that at least some individuals in the population could innovate foraging techniques and be more flexible, exploratory, and persistent, which could be learned by conspecifics and/or future generations.” The expectations about variance in addition to means are highly relevant.

RESPONSE 37: Great point, thank you for bringing this up! For the flexibility analysis, we now repeated the same simulation while holding the sample size constant, and setting all three site means to be the same and holding them constant, while we varied the standard deviation for each response variable in Q1. The results are in the new Table 3, and we added explanations about these results as follows:

Analysis Plan > Q1 > Flexibility Analysis: “To investigate the degree to which we can detect differences in the variances between sites, we ran another version of the **mathematical model** using a sample size of 15 per site and we held the mean number of trials to reverse a preference constant between all populations. We then changed the α standard deviations and performed pairwise site contrasts. We determined that it

will be difficult to detect meaningful differences in variances in the number of trials to reverse a preference between sites (Table 3).”

The results show that we will not be able to robustly detect differences in variance between populations because the boundary for where all of the values are on one side of zero moves around quite a lot. One thing that we have been discussing is the fact that measurement error can obscure differences in variances. Therefore, the simulation suggests that we will be able to detect differences in the mean with these sample sizes, but likely not differences in the variances.

For the other three analyses (innovation, exploration, and persistence), the distributions we used (binomial and gamma-Poisson) were such that the mean is tied to the variance, therefore, instead of attempting to pull variance out in these models, we will plot the variance for these variables to compare site differences visually. We believe this will be sufficient because the flexibility analysis showed us empirically that we will not be able to robustly detect differences between site variances, which is likely due to us choosing to model small sample sizes per site and that there is likely to be measurement error that, if large enough, can obscure differences in variance. We clarified this in the text as follows:

Analysis Plan > Q1 > Innovation Analysis: “Because the mean and the variance are linked in the binomial distribution, and because the variance simulations in the flexibility analysis showed that we will not be able to robustly detect differences in variance between sites, we will plot the variance in the number of loci solved between sites to determine whether the edge population has a wider or narrower spread than the other two populations.”

Analysis Plan > Q1 > Exploration Analysis: “Because the mean and the variance are linked in the gamma-Poisson distribution, and because the variance simulations in the flexibility analysis showed that we will not be able to robustly detect differences in variance between sites, we will plot the variance in the latency to approach the object between sites to determine whether the edge population has a wider or narrower spread than the other two populations.”

Analysis Plan > Q1 > Persistence Analysis: “Because the mean and the variance are linked in the binomial distribution, and because the variance simulations in the flexibility analysis showed that we will not be able to robustly detect differences in variance between sites, we will plot the variance in the proportion of trials participated in between sites to determine whether the edge population has a wider or narrower spread than the other two populations.”

COMMENT 38 (CN10): H1 > P1 alt 1 “If the original behaviors exhibited by this species happen to be suited to the uniformity of human-modified landscapes (e.g., urban, agricultural, etc. environments are modified in similar ways across Central and North America), then the averages and/or variances of these traits will be similar in the grackles sampled from populations across their range” This result may also occur if irrelevant behaviour have been tested, hence my comments above about ecological relevance of behavioural tests. Hence my concern about the ecological relevance of your experimental behavioural tests.

RESPONSE 38: We have now clarified the ecological relevance in the new Introduction.

COMMENT 39 (CN11): H1 > P1 alt 1 “Alternatively, it is possible that 2.9 generations at the edge site is too long after their original establishment date to detect differences in the averages and/or variances” It would be relevant to backup this by evidence from experimental evolution on learning skills in vertebrates (like mouse,...). I doubt that populations would get back to ancestral averages in cognition within 3 generations.

RESPONSE 39: We agree that we should back this statement up with evidence from experimental evolution - thank you for pointing this out! Evidence is accumulating that learning can be costly (reviews in Mery and Burns 2010 and Dunlap and Stephens 2016), and we found examples that we now include in the preregistration:

Prediction 1 > Alternative 1: “Alternatively, it is possible that 2.9 generations at the edge site is too long after their original establishment date to detect differences in the averages and/or variances (though evidence

from experimental evolution suggests that, even after 30 generations there is no change in certain behaviors when comparing domestic guinea pigs with 30 generations of wild-caught captive guinea pigs @kunzl2003wild, whereas artificial selection can induce changes in spatial ability in as little as two generations @kotschal2013artificial).”

Mery and Burns, 2010. Behavioral plasticity: an interaction between evolution and experience *Evolutionary Ecology*, 24 (2010), pp. 571-583

Dunlap and Stephens, 2016. Reliability, uncertainty, and costs in the evolution of animal learning *Curr. Opin. Behav. Sci.*, 12 (2016), pp. 73-79, 10.1016/j.cobeha.2016.09.010

COMMENT 40 (CN12): H1 > P1 alt 1 “If the sampled individuals had already been living at this location for long enough (or for their whole lives) to have learned what they need about this particular environment (e.g., there may no longer be evidence of increased flexibility/innovativeness/exploration/persistence), there may be no reason to maintain population diversity in these traits to continue to learn about this environment” Relevant : then focus on juveniles individuals in sampling, if possible

RESPONSE 40: We apologise for our lack of clarity. We focus on adult grackles for two key reasons: (i) they are more likely to have fully developed fine motor skills (e.g., holding/grasping objects with their bill – see Collias & Collias 1964 and Rutz et al. 2016 for ontogenetic differences in birds’ capacity to mandibulate nesting material and sticks, for example) and (ii) we cannot distinguish between, for example, a juvenile bird of 8 months versus an adult of 12 months of age. Thus, we do not focus on juvenile individuals so as not to confound potential age-related variation in cognitive abilities and in fine motor-skill development with variation in our target variables of interest. We now include this rationale:

Methods > Planned Sample: “Great-tailed grackles are caught in the wild in Woodland, California and at a site to be determined in Central America. We aim to bring adult grackles, rather than juveniles, temporarily into the aviaries for behavioral choice tests to avoid the potential confound of variation in cognitive development due to age, as well as potential variation in fine motor-skill development (e.g., holding/grasping objects—early-life experience plays a role in the development of both of these behaviors; e.g., Collias & Collias 1964, Rutz et al. 2016) with variation in our target variables of interest. Adults will be identified from their eye color, which changes from brown to yellow upon reaching adulthood (Johnson and Peer 2001).”

COMMENT 41 (CN13): Figure 2: For non-bird experts these pictures should be associated to explanations to justify the choice of behavioural tests.

RESPONSE 41: We added an ecological relevance statement for the choice of our behavioral tests in the new Introduction. However, we agree that our figure caption was too sparse, and so we added the below text. We changed two labels in our figure (“reversal learning” changed to “flexibility” and “multiaccess box” changed to “innovativeness”) to match the description:

Figure 2 > Legend: “Experimental protocol. Great-tailed grackles from core, middle, and edge populations will be tested for their: (top left) flexibility (number of trials to reverse a previously learned color tube-food association); (middle) innovativeness (number of options [lift, swing, pull, push] solved to obtain food from within a multi-access log); (bottom left) persistence (proportion of trials participated in during flexibility and innovativeness tests); and (far right) exploration (latency to approach/touch a novel object).”

COMMENT 42 (CN14): H2: “Changes in dispersal behavior, particularly for females, which is the sex that appears to be philopatric in the middle of the range expansion, facilitate the great-tailed grackle's geographic range expansion” Not clear to me why focus on females. Males is the usual dispersing sex. Females as the limiting factor (without females no nests)? Please clarify.

RESPONSE 42: We discovered that females are the philopatric sex in this species in a previous study (Sevchik et al. 2019 http://corinalogan.com/Preregistrations/gdispersal_manuscript.html), but, thanks to your comment, we realized this wasn't clear, therefore we added the citation. We also changed our predictions to make the expected effect clearer, in particular that we expect more dispersal at the edge. However, given

that we know that males disperse in the middle of the range expansion, we might only see an increase in dispersal at the edge for females:

Q2 > Prediction 2: “a higher proportion of individuals, particularly females, which is the sex that appears to be philopatric in the middle of the range expansion [@sevchik2019dispersal], disperse in a more recently established population”

COMMENT 43 (CN15): H2 > P2 “If a change in dispersal behavior is facilitating the expansion, then we predict more dispersal at the edge: a higher proportion of individuals disperse in a more recently established population and, accordingly, fewer individuals are closely related to each other” This appears to be true in many species, but it may be necessary but not sufficient to colonize new areas. Innovation may be needed in addition to increased dispersal at distribution edges.

RESPONSE 43: We now made it clearer throughout the text that we are not talking about competing alternative hypotheses, but that the range expansion could be associated with all or none of the variables we are measuring (see Response 4). For example, we might find that the individuals in the edge population show higher levels of innovation as well as being more likely to have dispersed and we cannot and do not intend to tease these apart. We made sure to clarify this as follows:

Q2 > P2: “We predict **more dispersal at the edge**: a higher proportion of individuals, particularly females, which is the sex that appears to be philopatric in the middle of the range expansion [@sevchik2019dispersal], disperse in a more recently established population and, accordingly, fewer individuals are closely related to each other. *This would support the hypothesis* that changes in dispersal behavior **are involved in** the great-tailed grackle's geographic range expansion.”

COMMENT 44 (CN16): H2 > P2 alt 1 “If the original dispersal behavior was already well adapted to facilitate a range expansion, we predict that the proportion of individuals dispersing is not related to when the population established at a particular site and, accordingly, the average relatedness is similar across populations.” This explains that relatedness measures are collected. However do you have evidence that your markers for quantifying relatedness (microsat?! got it is snp later on) will be variable enough to detect such limited changes in relatedness? Would another behavioural test perhaps be a better estimate of dispersal (perhaps, propensity to leave a cage for another in large field enclosures, perhaps?)?

RESPONSE 44: Thanks to your comment, we realized this wasn't clear. We have evidence that these markers work for quantifying relatedness because we have already conducted a dispersal study on part of the Arizona great-tailed grackle population (Sevchik et al. 2019 http://corinalogan.com/Preregistrations/gdispersal_manuscript.html). We now explain this in better detail in the new Introduction:

Introduction: “To determine whether females and/or males move away from the location they hatched, we will assess whether their average relatedness (calculated using single nucleotide polymorphisms, SNPs) is lower than what we would expect if individuals move randomly [@sevchik2019dispersal].”

COMMENT 45 (CN17): Table 1: “The number of generations at a site is based on a generation length of 5.6 years for this species (@GTGRbirdlife2018) and on the first year in which this species was reported to breed at the location” At what age do they start to breed ?

RESPONSE 45: They start breeding at age 1. We added this to:

Table 1 legend: “(note: this species starts breeding at age 1)”

COMMENT 46 (CN18): Table 1: Nice contrasted population sites

RESPONSE 46: Thank you!

COMMENT 47 (CN19): H3 > P4 “Over the past few decades, GTGR has increased the habitat breadth that they can occupy, whereas BTGR continues to use the same limited habitat types.” Re: Habitat breadth: Which are ?

RESPONSE 47: We now include known habitat differences between these two species in the new Introduction, which appear to be related to suitable nesting habitat - thank you for pointing out that we did not include this information! This is what we added:

Introduction: “Detailed reports (@selander1961analysis, @wehtje2003range) on the breeding ecology of these two species indicate that range expansion in boat- but not great-tailed grackles may be constrained by the availability of suitable nesting sites. Boat-tailed grackles nest primarily in coastal marshes, whereas great-tailed grackles nest in a variety of locations (e.g., palm trees, bamboo stalks, riparian vegetation, pines, oaks). However, this apparent difference in habitat breadth has yet to be rigorously quantified.”

COMMENT 48 (CN20): H3 > P5 “Some inherent trait allows GTGR to expand even though both species have unused habitat available to them.” It would be relevant to quantify behavioural traits in the sister species as well.

RESPONSE 48: We agree and we have future plans to do a behavioral comparison with BTGR, however it is beyond the scope of our current funding period.

COMMENT 49 (CN21): Figure 3 “Comparing the availability of suitable habitat between great-tailed grackles (GTGR), which are rapidly expanding their geographic range, and boat-tailed grackles (BTGR), which are not” They certainly have different habitat requirements given that their distribution ranges do not overlap. It will be hard to make a useful comparison between the two species without quantifying behavioural traits in the sister, not expanding, species.

RESPONSE 49: Please see our Response 48. The ranges of the two species do overlap in Texas, Louisiana, Mississippi, and Alabama (Selander & Giller 1961, eBird.org). Additionally, because our hypotheses about behavior and habitat are not mutually exclusive, we can still determine whether habitat changes play a role in the boat-tailed grackle’s lack of a rapid range expansion.

COMMENT 50 (CN22): Methods > Planned sample “Great-tailed grackles are caught in the wild in Woodland, California and at a site to be determined in Central America” Focus on juveniles (as suggested above)?

RESPONSE 50: Please see our Response 40.

COMMENT 51 (CN23): Methods > Planned sample: “We catch grackles with a variety of methods (e.g., walk-in traps, mist nets, bow nets), some of which decrease the likelihood of a selection bias for exploratory and bold individuals because grackles cannot see the traps (i.e., mist nets)” good

RESPONSE 51: Thank you!

COMMENT 52 (CN24): Methods > Data collection stopping rule: “We will stop collecting data on wild-caught grackles in H1 and H2 (data for H3 are collected from the literature)” this is very surprising : what type of data ? please specify.

RESPONSE 52: We now clarified this in the new Introduction:

“Secondly, we aim to investigate whether habitat availability, not necessarily inherent species differences, explains why great-tailed grackles are able to much more rapidly expand their range than their closest relative, boat-tailed grackles (*Q. major*) [@post1996boat; @wehtje2003range]. Detailed reports on the breeding ecology of these two species indicate that range expansion in boat- but not great-tailed grackles may be constrained by the availability of suitable nesting sites [@selander1961analysis; @wehtje2003range]. Boat-tailed grackles nest primarily in coastal marshes, whereas great-tailed grackles nest in a variety of locations (e.g., palm trees, bamboo stalks, riparian vegetation, pines, oaks). However, this apparent difference in habitat breadth has yet to be rigorously quantified. Great-tailed grackles inhabit a wide variety of habitats (but not forests) at a variety of elevations (0-2134m), while remaining near water bodies, while boat-tailed grackles exist mainly in coastal areas [@selander1961analysis]. Both species have similar foraging habits: they are generalists and forage in a variety of substrates on a variety of different food items [@selander1961analysis]. We will use ecological niche modeling to examine temporal habitat changes over

the past few decades using observation data for both grackle species from existing citizen science databases. We will compare this data with existing data on a variety of habitat variables. We identified suitable habitat variables from @selander1961analysis, @johnson2001great, and @post1996boat (e.g., types of suitable land cover including marine coastal, wetlands, arable land, grassland, mangrove, urban), and we added additional variables relevant to our hypotheses (e.g., distance to nearest uninhabited suitable habitat patch to the north, presence/absence of water in the area). A suitable habitat map will be generated across the Americas using ecological niche models. This will allow us to determine whether the range of great-tailed grackles, but not boat-tailed grackles, might have increased because their habitat suitability and connectivity (which combined determines whether habitat is available) has increased, or whether great-tailed grackles now occupy a larger proportion of habitat that was previously available.”

COMMENT 53 (CN25): Methods > Protocols and open materials > Suitable habitat: “We identified suitable habitat variables from Selander and Giller (1961), Johnson and Peer (2001), and Post et al. (1996), and we added additional variables relevant to our hypotheses. A suitable habitat map will be generated across the Americas using GIS. ” This is central to explain because it is not straightforward to see the relevance and feasibility

RESPONSE 53: Please see the new Introduction where we clarified this.

COMMENT 54 (CN26): Analysis plan: This seems very well done but I have not read with total attention

RESPONSE 54: Thank you very much! It was our first time fully implementing what we have been learning from Richard McElreath’s Statistical Rethinking book and course. We’re really proud that we were able to develop these models because they are much better suited to our questions than standard analyses.

COMMENT 55 (CN27): Analysis Plan > P3 > Explanatory variables 1-6. Here are the variables for habitat comparison. It would be useful to specify what range of values for these variables are relevant for each of the two species (is it known?)

RESPONSE 55: The range of values for these two species is not known, and it is one of the aims of our investigation to determine this. Please see our Response 29 for many more details on each explanatory variable.

To help clarify, we added to the Analysis Plan > Q3 which analyses we will run to answer our questions:

Analysis 1 (P3: different habitats): does the range of variables that characterize suitable habitat for GTGR differ from that of BTGR? We will fit species distribution models for both species in 2018 to identify the variables that characterize suitable habitat. We will examine the raw distributions of these variables from known grackle occurrence points or extract information on how the predicted probability of grackle presence changes across the ranges for each habitat variable. The habitat variables for each species will be visualized in a figure that shows the ranges of each variable and how much the ranges of the variables overlap between the two species or not.

Analysis 2 (P3: habitat suitability): has the available habitat for both species increased over time? We will fit species distribution models for both species in 1970 and in 2018 and determine for each variable, the range in which grackles are present (we define this as the habitat suitability for each species). Then we will take these variables and identify which locations in the Americas fall within the grackle-suitable ranges in 1970 and in 2018. We will then be able to compare the maps (1970 and 2018) to determine whether the amount of suitable habitat has increased or decreased.

If we are able to find data for these variables before 1970 across the Americas, we will additionally run models using the oldest available data to estimate the range of suitable habitat earlier in their range expansion.

Analysis 3 (P3: habitat connectivity): has the habitat connectivity for both species increased over time? If the connectivity distances are smaller in 2018, this will indicate that habitat connectivity has increased over time. We will calculate the least cost path from the northern edge to the nearest suitable habitat patch. To

compare the distances between 1970 and 2018, and between the two species, we will run two models where both have the distance as the response variable and a random effect of location to match the location points over time. The explanatory variable for model 1 will be the year (1970, 2018), and for model 2 it will be the species (GTGR, BTGR).

If we are able to find data for these variables before 1970 across the Americas, we will additionally run models using the oldest available data to estimate the range of connected habitat earlier in their range expansion.

Analysis 4 (P4: habitat breadth): has the habitat breadth of both species changed over time? We will count the number of different land cover categories each species is or was present in for 1970 and 2018. To determine whether this influences their distributions, we will calculate how much area in the Americas is in each land cover category, which would then indicate how much habitat is suitable (based solely on land cover) for each species.