

Dear Dr.'s Vogel, Gingins, and anonymous reviewers,

We sincerely apologize for the delay in our revision. Due to some staffing changes that occurred in the past few months, all of us were overcommitted just by trying to keep the experiments and field site running. Logan was in the field collecting data to help offset the setbacks, which meant that she was unable to lead the revision process until now.

Luckily, we received your reviews before more data were collected so we were able to determine whether any methodological changes were needed in advance, which is the point of the preregistration process.

One general note: as a result of several setbacks, we had to break the counterbalanced order of experiments for four birds in batch 1. For details, please see Table 1 in the protocol: https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwcKOB_KcT7HE/edit?usp=sharing.

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. We have revised our preregistration (https://github.com/corinalogan/grackles/blob/master/EasyToReadFiles/g_inhibition.md) and protocol (https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwcKOB_KcT7HE/edit?usp=sharing), and we responded to your comments (which we numbered for clarity) below (our responses are preceded by "Response X").

We think the revised version is much improved due to your generous feedback!

All our best,
Corina, Kelsey, Zoe, Luisa, Benjamin, Aaron, and Claudia

Are the more flexible great-tailed grackles also better at inhibition?

Corina Logan, Kelsey McCune, Zoe Johnson-Ulrich, Luisa Bergeron, Carolyn Rowney, Benjamin Seitz, Aaron Blaisdell, Claudia Wascher

10.17605/OSF.IO/GCA5V version v1.4

Submitted by Corina Logan 2018-10-12 18:36

Abstract

This is a PREREGISTRATION. The DOI was issued by OSF and refers to the whole GitHub repository, which contains multiple files. The specific file we are submitting is `ginhibition.Rmd`, which is easily accessible at GitHub at <https://github.com/corinalogan/grackles/blob/master/ginhibition.Rmd> (note: the PCI website tends to delete underscores, which breaks this link. There is an underscore between "g" and "inhibition.Rmd"). Viewing this file at OSF will result in NOT being able to see the figure as part of the .Rmd file (use the GitHub link instead). Photo credit: Corina Logan (CC-BY-SA 4.0). We will likely start data collection in late November 2018 so it would be ideal if we could get through the review process before then.

Keywords: Behavioral flexibility, inhibition, self control, delay of gratification, go no-go, detour, comparative cognition, avian cognition
Round #1

Decision

by Erin Vogel, 2018-12-09 23:34

Manuscript: 10.17605/OSF.IO/GCA5V

Revise and resubmit

Dear Dr. Logan,

I apologize on the delay for the reviews on your preregistration. We have now received three reviews (attached below) and I have read over the reviews and your preregistration. Overall, your preregistration was excellent and the reviewers agreed with minor revisions it should be accepted. Thus, my recommendation is to address these minor points that will greatly clarify your proposal and resubmit. Once these points are addressed, the recommendation will be approved.

See attached PDF for comments. If you have any additional comments, please contact me.

Best wishes,

Erin Vogel

Reviewed by anonymous reviewer, 2018-11-19 07:54

1. This is an interesting study that examines the relationship between behavioural flexibility and inhibition using great-tailed grackles as a study model. This investigation for the understanding about whether flexibility as a trait would co-evolve with other traits in facilitating a species establishment and expansion. Overall, I find the study design is appropriate; tasks are set to measure behavioural flexibility as well as inhibition. However, as the submitted document is a preprint (as opposed to a written manuscript with completed data collection and analyses), my comments here are solely down for the submitted version. I hope the authors could address each comment during the revision.

Response 1. Thank you so much for your feedback! We are still learning how to format these preregistrations to make it easier for us in that we don't have to write the final papers in advance, but where we give enough detail so reviewers have enough context to evaluate the planned research. We really value your input on how we can do better with this.

2) Rationale of the study: The title itself highlights the investigation interest lies in behavioural flexibility and inhibitory control. Yet, the abstract and the introduction have not even explained why inhibitory control would be selected as the main investigation of the study. Accordingly, some information about the relationship between behavioural flexibility and how it is likely related to other traits, as well as the reasons for selecting inhibition in particular, would have strengthen the rationale for the study. This additional information will greatly help to work out the logics in hypotheses testing.

Response 2. Excellent point. We made the following changes:

Abstract: “In this piece of the long-term project, we aim to test whether the assumption that inhibition is required for flexibility (which recent brain scanning and genetic studies do not support) by measuring grackle inhibition in three widely used experimental paradigms (delay of gratification, go-no go, detour) assessing different aspects of inhibitory control to determine whether those individuals that are more flexible are also better at inhibiting. This species likely uses inhibition in the wild by inhibiting behavioral responses to food sources (e.g., we have seen them wait to steal food from a human's plate until the human turned away from the table) and potentially modulating its responses in the presence of dominant individuals (e.g., sneaker males may inhibit copulating with a female until the territorial male is absent).”

Hypothesis: “A common assumption is that inhibition is required for an individual to exhibit behavioral flexibility (e.g., Manrique et al. 2013, Griffin & Guez 2014, Liu et al. 2016), however brain scanning and genetic evidence suggests this is not the case (e.g., Homberg et al. 2007, Ghahremani et al. 2010).”

3) Rationale of the tasks: Authors should highlight the reason(s) for using the three inhibition tasks when measuring inhibition, in particular there are some critics about the use of the cylinder task (van Horik et al., 2018).

Response 3. We agree that there are substantial issues with using the cylinder task as a measure of inhibition, however we also acknowledge that the test is widely used in comparative cognition studies, often as a singular measure to assess inhibitory control in a given species (e.g., Isaksson et al. 2018, Langbein 2018, Can et al. 2016, Bobrowicz & Osvath 2018). We believe more comparative studies, assessing correlations in performance in different inhibitory task paradigms (e.g., Brucks et al 2017) are urgently needed. This is the reason we wanted to include it in our test battery: to see if it is actually a measure of inhibition, which we think the other two tests (go no-go and delay of gratification) actually measure. If detour performances correlate positively with the other two tests, then it would provide some evidence that it might measure an inhibitory response.

We have revised the preregistration to clarify this by adding to the abstract: “we aim to measure grackle inhibition in three widely used experimental paradigms (delay of gratification, go-no go, detour) assessing different aspects of inhibitory control”

And we added to Hypothesis > P2 alternative:

“Note that we consider these two tasks as valid measures of inhibition, whereas it is questionable whether the detour task actually measures inhibition (e.g., Van Horik et al. 2018). If the detour results correlate with performances on the other two inhibition tasks, this would provide some evidence that the detour task is a measure of inhibition.”

Isaksson, E., Utku Urhan, A., & Brodin, A. (2018). High level of self-control ability in a small passerine bird. *Behavioral Ecology and Sociobiology*, 72(7), 118.
<http://doi.org/10.1007/s00265-018-2529-z>

Langbein J. 2018. Motor self-regulation in goats (*Capra aegagrus hircus*) in a detour-reaching task. PeerJ 6:e5139 <https://doi.org/10.7717/peerj.5139>

Can, K., A., T. L., P., von B. A. M., & Mathias, O. (2016). Ravens, New Caledonian crows and jackdaws parallel great apes in motor self-regulation despite smaller brains. *Royal Society Open Science*, 3(4), 160104. <http://doi.org/10.1098/rsos.160104>

Bobrowicz, K., & Osvath, M. (2018). Cats Parallel Great Apes and Corvids in Motor Self-Regulation – Not Brain but Material Size Matters . *Frontiers in Psychology* . Retrieved from <https://www.frontiersin.org/article/10.3389/fpsyg.2018.01995>

Brucks, D., Marshall-Pescini, S., Wallis, L. J., Huber, L., & Range, F. (2017). Measures of Dogs' Inhibitory Control Abilities Do Not Correlate across Tasks . *Frontiers in Psychology* . Retrieved from <https://www.frontiersin.org/article/10.3389/fpsyg.2017.00849>

4) Clarity in concepts: Authors defined behavioural flexibility is the behavioural changes with increased experience or the outcome of learning. While flexibility in discrimination-reversal learning task is clearly learning based, flexibility in novel food-extraction task (or some forms of variants as in multi-access box in this project) has been discussed is a different form of flexibility from those that seen in discrimination-reversal learning task (Audet & Lefebvre, 2017). In this case, it is not entirely sure what forms of flexibility authors are measuring and if these forms of flexibility are correlated with inhibition at all. Inhibition in novel food-extraction tasks could be seen in a design that requires individuals to absolute a previously learned technique in the face of a similar task. At any rate, could authors clarify which form of flexibility they would like to measure?

Response 4. There are many, many definitions of behavioral flexibility and we agree that it is important to define what we mean by it. Our definition of behavioral flexibility is discussed in detail in Mikhalevich et al. 2017, but perhaps that wasn't quite clear. Your comment about how novelty plays into our definition of flexibility is a good one so we clarified what we mean. Specifically, our definition of flexibility is about switching between options, and the learning from previous experience can come not only from what they learned about this particular task, but also what they have learned in the wild in general about human-made apparatuses that they are trying to break into to get food. We added:

Hypothesis: "see Mikhalevich et al. 2017 for details"

Hypothesis: "as measured by reversal learning (where they must learn to prefer one of two options that contain food and then reverse this preference) and switching between options on a multi-access box (where they must learn to switch to a new option, out of four available options, when an option becomes non-functional). We expect this species to be behaviorally flexible because they are fast at reversal learning (Logan 2016), they often encounter human-made "puzzle boxes" in the wild as they attempt to open packaging to access food when digging through garbage cans and eating at outdoor cafes, and they may track resources across time and space."

5) Procedures of tasks: a) P.3 Figure 2. Add 'but far from the bird' in '(near the experimenter)'; this will make a contrast statement for '(near the bird)'.

Response 5. Thank you for the suggestion! We ended up needing to modify the delay of gratification apparatus to better work for the grackles who are not hand-raised and therefore would be uncomfortable with an experimenter's hand entering their aviary. We made an apparatus that has three cups, each with its own lid that the experimenter is able to pull open from the aisle, therefore making food available after a delay. Please see our revised protocol for details and pictures (https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwckOB_KcT7HE/edit?usp=sharing).

6 b) P.3 Figure 2. Three trials, consecutive trials?

Response 6. Good question, we hadn't yet decided this. We now updated the protocol (page 9, https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwckOB_KcT7HE/edit?usp=sharing) to say three trials in one session.

7 c) P.3 Figure 2. In the sentence 'Once they pass training (by waiting for more than one food item in three trials), they move on to the test where food items are transferred from the serving to the storing lid with delays ranging from 2-1280 seconds.', 'food items are transferred from the serving to the storing lid' should be 'from the storing to the severing lids', right?

Response 7. Yes you are correct, thank you for catching this! We ended up changing this text to accommodate the new apparatus we made for the grackles (see Response 5).

8 d) Colour discrimination test - I understand authors have counterbalanced the colour presentation for birds in the colour discrimination test. However, are the chosen green and purple colours neutral to the birds (i.e. birds should not show a colour preference in either colour)? This is because if a bird shows any preference to one colour and that colour is rewarded in the discrimination phase, it could mask the true learning rate. If a bird is rewarded with the non-preferred colour in the discrimination phase, this would be measuring inhibition.

Response 8. It turns out that in the process of training the first grackles to use the touchscreen, we discovered that they are scared of either colors on the screen, or circles that might look like eyes. Therefore, we modified our stimuli to be white with gray lines (therefore they differ in pattern, having fewer or more lines) and to be shapes other than circles. What was to be the pink stimulus is now a white rectangle with fewer horizontal lines, and what was to be the green stimulus is now a white rectangle with more horizontal lines (see pictures in the protocol:

https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwcKOB_KcT7HE/edit?usp=sharing). This also has the benefit of addressing your valid comment about potential color preferences influencing learning rates.

9 e) To speed up birds' learning process in using touch screen, authors could also consider shaping. The stimulus (0.5cm diameter) in the 'moving dot' phase might be relatively small for birds, which would reduce the probability for birds to hit the target and likely prolong the training process. Therefore, shaping process in which a much larger stimulus (say 2.5cm diameter) at this stage would increase the probability of pecking the touch screen for a bird, followed by slowly reducing the diameter of the stimulus to 0.5cm may facilitate birds' learning to use touch screen.

Response 9. Thank you so much for your ideas on how to train the grackles to use a touchscreen! Indeed, the process for training them has been really different from what some of us (Aaron and Benjamin) have previously experienced with pigeons, so we value all ideas here. It turns out that the grackles didn't attend to the moving dot at all and what ended up working was just going straight to the white square food key training program and hand shaping them to peck the white square (at first drawing their attention to the digital square by taping a white piece of paper to the screen until they transferred to the digital square). We have updated our training methods in the protocol (https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwcKOB_KcT7HE/edit?usp=sharing).

10 f) Authors might also want to consider a habituation phase to the touch screen for birds to explore the testing environment/operant box or, to minimise any neophobic responses. Also, make sure birds are always tested in the same compartment/box.

Response 10. Good call! You are so right about this - they need lots of habituation to the touchscreen apparatus so what we do is put the whole thing in their aviary and feed them out of the food hopper when they are not being tested. It generally takes at least a few days for them to become comfortable. Good point about always using the same touchscreen for the same birds. We only have one touchscreen so that won't be a problem. We added this to the protocol (page 14 https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwcKOB_KcT7HE/edit?usp=sharing)

Reviewed by anonymous reviewer, 2018-12-03 17:17

11. This is an interesting proposal which has the potential to answer very important questions about a topic critical to behavioral ecology – the role that inhibition may (or may not) play in the evolution of behavioral flexibility.

Response 11. Thank you very much! We are glad you think this will be a useful contribution.

12. The authors provide a brief introduction to the project, focused hypotheses/ predictions, and great detail about their project timeline and methods. My comments for each section are below:

Response 12. Thank you so much for your comments!

Abstract.

13. o The background makes it seem like the goal of this project will be to test if our test-based measures of behavioral flexibility can reliably predict realized behavioral flexibility (i.e. answering the question – can we use tests of behavioral flexibility in order to predict a species' ability to move into a new environment)? However, the goal of this project is to test whether or not behavioral flexibility predicts inhibition and the consistency of relevant tests. I suggest restructuring the abstract to reflect the goals of this project – they need to explicitly address why linking behavioral flexibility to inhibition is interesting and important to their overall question.

Response 13. Yes, good point, we needed to tie in inhibition more directly. Please see response 2 for details on how we addressed your comment.

Predictions.

14. o P1 is well structured, as are the alternatives. However, if the hypothesis is that “flexibility requires inhibition”, then they should test whether or not inhibition predicts flexibility (rather than the other way around).

Response 14. Good point - the flexibility measures are listed as independent variables so we had already set it up this way, but we missed changing the language in the Hypothesis section. We made the Hypothesis P1 text consistent.

15. o P2 does not follow from the hypothesis – I suggest switching P2 with its alternative to maintain a consistent structure.

Response 15. Good call - we made the change.

16. o The authors should explain why the go-no go test is being validated against the delayed gratification test (rather than against the detour test or vice versa).

Response 16. Inhibitory control is a multidimensional construct incorporating aspects such as motor-inhibition and the ability to delay gratification. A correlation between different aspects of inhibitory control has been shown in humans (Duckworth & Kern 2011), however the relationship or independence of these different cognitive aspects in non-human animals is not very well understood and therefore we feel the evaluation of such a relationship is of great interest. Regarding the choice of experimental paradigms, we have chosen three paradigms widely applied in human as well as non-human literature. Detour tasks are very

popular in comparative cognition and applied in a wide variety of species, however we also share some concerns regarding the validity of the obtained results (please see our response 3 above). We therefore seek to evaluate our findings comparing the performance in the go-no go test and the delay of gratification experiment because these two tests are presumed to measure the same type of inhibitory control. We will further determine whether performances in the go-no go test correlate with performances in the detour task, which would be strongly expected if both tasks measure motor-inhibition. We made the following changes to address this:

Hypothesis > P2: “If go no-go task performance strongly correlates with performance on the delayed gratification task, both of which measure the same type of inhibition, this indicates these two tasks measure the same trait, which therefore validates an inhibition task using a touch screen (the go no-go task). Note that the detour task measures a different type of inhibition (and it is questionable whether it actually measures inhibition; e.g., Van Horik et al. 2018), which is why we are validating the go no-go task against the delayed gratification task. If the detour results correlate with performances on the other two inhibition tasks, this would provide some evidence that the detour task is a measure of inhibition as is claimed.”

Analysis plan > P2: we added an analysis to examine whether go no-go results correlate with detour performances, and an analyses to determine whether detour and delay of gratification performances correlate.

Duckworth, A. L., & Kern, M. L. (2011). A meta-analysis of the convergent validity of self-control measures. *Journal of Research in Personality*, 45(3), 259–268.
<https://doi.org/https://doi.org/10.1016/j.jrp.2011.02.004>

17. o P3 needs to be restructured to follow that of P1 – include P3 and alternatives separately.

Response 17. Nice catch! We separated out the negative correlation into the new P3 alternative. In terms of following the structure of P1, we are not quite sure what you mean because this prediction is asking a different question from that in P1.

Methods.

18. o As written, the methods and analysis plan are difficult to follow. I suggest organizing the methods per prediction (i.e. listing the dependent/independent variables under each section).

Response 18. This is an excellent suggestion to which we developed an alternative solution - we placed tags throughout the preregistration so the reader can click and jump to the next section for that prediction. For example, in the Hypothesis section, at the end of prediction 1, we have a tag that allows the reader to jump to prediction 1’s dependent variables, and from there to the P1’s independent variables, and from there to the P1 analyses.

Reviewed by Simon Gingins, 2018-11-20 12:34

General comments:

19. This preregistration describes a series of experiments (detour, go no-go, delayed gratification) in order to investigate inhibition in the great-tailed grackle. The goal is then to combine these results with other experiments on flexibility on the same individuals, and ask whether behavioural flexibility requires a certain level of inhibition. While the role of inhibition in cognition has been broadly studied, few or no studies have attempted to understand the relationship between inhibition and flexibility. The experiments described are well designed, the predictions clearly laid out and the analysis appears to be statistically sound. Therefore, I believe that this project would constitute a welcome addition to the literature once it is completed.

Response 19. Thank you so much for your positive feedback! We are so glad you think this will be a worthy contribution to the literature.

20. Nevertheless, I have a few reservations with regards to the preregistration itself. My main comment is that it is hard for the reader to fully understand the rationale and the execution of the experiments. This project is part of a larger scale endeavour, and little effort was made to make all the relevant information easily accessible in this preregistration, sometimes making it difficult to assess. Additionally, the structure is confusing at times. For instance, all the details of the methodology are given in the very last section, while many of the predictions and other aspects of the experimentation are described elsewhere. It would be easier to understand if the rationale, the methodology, the predictions and the analysis for each experiment was given as a clear sequence.

Response 20. Thank you for your feedback on how we can do better with the preregistration process! One of the difficulties we have run into is that, for example, the flexibility preregistration (https://github.com/corinalogan/grackles/blob/master/EasyToReadFiles/g_flexmanip.md) is undergoing its own peer review process at PCI Ecology and we wanted to make it clear to reviewers of each of the preregistrations what it was they were being asked to review. This is why we preferred to keep them separate from each other in their own files and provide links to the separate preregistrations so people could see the project in “units”. We moved Planned Sample to the beginning of the Methods section to accommodate the more logical flow of information. Please see response 13 for details on how we integrated the inhibition tests into the broader framework more clearly. Please see response 18 for details on how we made the preregistration easier to navigate when the reader wants to look at one prediction at a time.

21. My second main comment is that it is unclear why great-tailed grackles are a good system for answering the questions of interest. The main reason given to study flexibility & inhibition in this system is that “they have rapidly expanded their range into North America over the past 140 years”. However, little information is given on the ecology of this species. Furthermore, I feel that ecological relevance in the cognitive tasks has been neglected. Species-specific traits can promote or impair individuals’ performance in laboratory tasks,

and hence I believe it is important to integrate the ecological challenges encountered by these birds when predicting the outcome of experiments. For instance, under what circumstances is inhibition beneficial for these birds? Do they require to be behaviourally flexible in their expanded range? Do they often need to choose between several options? Is there a lot of competition between individuals for accessing food? What do they typically feed on? How do they access this food? Such attributes might affect the outcome of experiments. It would thus be wise to predict, in the light of the grackles' ecology, which experiments might represent an additional challenge, and which ones might be more straightforward for them.

Response 21. We very much agree that it is important to bring ecological relevance to comparative cognition tests so we thank you for your prompting! We added the following to the preregistration:

Hypothesis: "We expect this species to be behaviorally flexible because they are fast at reversal learning (Logan 2016) and they may track resources across time and space. For example, we have observed them attending to when restaurants with outdoor cafes are particularly busy, and we noticed them change their restaurant preferences after a restaurant permanently closes. This species also likely uses inhibition by inhibiting behavioral responses to food sources (e.g., we have seen them wait to steal food from a human's plate until the human turned away from the table) and potentially modulating its responses in the presence of dominant individuals (e.g., sneaker males may inhibit copulating with a female until the territorial male is absent)."

P1: "An example of when a grackle might need to use both flexibility and inhibition in the wild is if a sneaker male keeps track of which females are on which male's territories and then chooses to copulate with females he has not previously copulated with when it is most likely that the territorial male is not present."

P2 alternative: "For example, if their performance is poor on the go no-go task, but better on the delay of gratification task, this might mean that the touchscreen is not an ecologically relevant enough context in which to exercise proactive inhibition, whereas the delay of gratification apparatus might be more similar to something they have encountered in the wild (e.g., we have seen grackles wait until French fries are dumped into a garbage can before diving into the garbage can to retrieve them)."

P3: "(and also potentially in the wild because they often lift up objects to look for food underneath)"

Please also see our response 2 for details on how we added ecological relevance to the abstract.

Specific comments:

Please note that my comments refer to the pages of the PDF I was given for review.

Page 3

Section C

22. P1: Why test for higher quantity of food rewards if this is not linked to inhibition? Does it bring additional information? Or is it to confirm the results from Hillemann et al. 2014?

Response 22. In order to avoid confusion, we revised the sentence and removed the reference to quantity at this point. The delay of gratification test involves a decision making process by the individual and previously it has been shown in corvids that although they were willing to wait for a higher quality reward, they did not do so for a higher quantity. In our experiment, we aim to replicate these previous results, hence testing the grackles in the quality and quantity condition.

23. P1 alternatives: It seems that it will be virtually impossible to disentangle between these two alternatives. How will this be interpreted?

Response 23. We agree with you. We just wanted to be clear in our preregistration all of the ways we might interpret our results (in the discussion sections of the future papers) after we have the data. If there is no correlation between flexibility and inhibition measures, we will interpret this as alternatives 1 and 2 both being possible and we will suggest that future research is designed to resolve this issue.

24. P2 alternative: what about the detour task?

Response 24. Please see our response 16.

Figure 2.

25. "...where food items are transferred from the serving lid to the storing lid with delays ranging from 2-1280 seconds"

Response 25. Sorry, we are a bit unclear about what the suggestion is. Is the suggestion to replace the text in Figure 2 Inhibition 2. Test

"Items transferred with delay: 2, 5, 10, 20, 40, 60, 80, 160, 320, 640, 1280s"

With

"where food items are transferred from the serving lid to the storing lid with delays ranging from 2-1280 seconds"?

26. Shouldn't it be "from the storing lid to the serving lid"?

Response 26. Yes, sorry about that. We ended up changing this text completely due to modifying the apparatus to meet the needs of this species. Please see our response 5 for more details.

27. Go no-go task: why introduce the negative stimulus? For each coloured dot there is a correct & an incorrect behaviour, one could thus simply reward all correct behaviour, equally for both situations. I am wondering whether the introduction of a negative stimulus will actually help understand the results. In both cases, the bird must choose to peck or to refrain to peck. When it pecks, it can get a reward. When it doesn't peck, it can only avoid getting the negative stimulus, but never gets a reward. A simple strategy could be to just peck all the time, and have exactly the same amount of rewards than if it did the task correctly, although with additional delays. I fear that it will remain unclear how the negative stimulus affected the bird's ability to solve this task. If the negative stimulus has a huge impact, birds might learn fast, but if this is a minor disturbance, they might accept many more errors. Similarly, is it mostly the stimulus itself, or the added delay that is influencing learning? I believe it would be more straightforward to give rewards for correct choices, and no rewards for incorrect choices, all the time.

Response 27. Thank you for the feedback. It was Logan's idea to use the aversive stimuli for incorrect choices because it seemed that there wouldn't be much of a cost to making the incorrect choice on the non-rewarded stimulus. However, we now know more about how grackles interact with touchscreens (from touchscreen training and a reversal learning experiment on the touchscreen) and it turns out that they are much more afraid of particular stimuli than we previously expected. Therefore, we removed the aversive visual and auditory stimuli from this experiment. Instead, if they peck the food key on unrewarded trials (when the correct response is to inhibit pecking), nothing will happen, no food reward will be available, and they will need to wait for the next trial to begin.

Page 5

28. P3: why only two and not 1/2? Testing half of the individuals would allow for more power in the comparisons, and this would provide a fully counterbalanced design. Furthermore, what will happen to these two birds doing the detour task before the flexibility manipulations, will they also experience the other inhibition manipulations (delayed gratification, go no-go)? If so in which order?

Response 28. We chose only two birds per batch to experience the detour task before the reversal learning tube experiment because this is a side question that we want to test. The main question we are interested in is whether manipulating flexibility has an effect on other traits such as inhibition. Keeping the bulk of the individuals receiving detour after reversal tubes gives us more statistical power to answer our main question. We now note this in Table 1 in the protocol. The birds that get detour before reversal tubes experience delayed gratification and go no-go in counterbalanced order, which, thanks to your comment, made us realize that we hadn't updated Table 1 to reflect this. It is now up to date: https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwcKOB_KcT7HE/edit?usp=sharing.

Page 6

29. P1.2.a 85% correct over how many trials? I now see this is described at the end of the manuscript. Would be good to include this here, or at least to refer to the section where the details are explained.

Response 29. Good call! We clarified the text in the preregistration to match what we wrote in the protocol.

Page 7

30. Flexibility 1 & 2: What is “the last reversal an individual experienced”? Similarly, it is not clear what is “the first 40 trials in their final reversal after the individual has seen the newly rewarded option once”. I understand this must be described in another preregistration, but the reader is left with the task of reading a whole other preregistration to figure out what is going on. Maybe expand this section a little to give more details? I think just a couple sentences might be enough. Or at least point to the relevant section of the other preregistration?

Response 30. Sorry for the confusion! We now provide explanations of what we mean. Independent variables > P1: “Flexibility 1: Number of trials to reverse a preference in the last reversal (in the reversal learning experiment) an individual experienced (individuals in the flexibility control group only experience 1 reversal so this data will come from their first and only reversal; individuals in the flexibility manipulation group experience serial reversals until they pass a certain criterion, therefore we will only use data from their most recent reversal)”

Independent variables > P1: “Flexibility 2: The ratio of correct divided by incorrect trials for the first 40 trials in their most recent reversal after the individual has seen the newly rewarded option once (i.e., they have explored the two color options enough to have discovered which color the food is now associated with). We use the first 40 trials because this should make it so that all subjects the same amount of data to compare. So far, all grackles require them more than 40 trials to pass the criterion to indicate they have reversed a preference.”

31. It is quite unclear what is the measure described in “flexibility 4”.

Response 31. This is a measure we just started developing with Richard McElreath. It is very much a work in progress and at this point we don’t know much more about it, but we will finish developing the measure before we conduct any of the analyses in the Analysis Plan. Sorry we can’t be more concrete yet - we just wanted to make sure we accounted for the overall goal of this measure in the preregistration from the beginning.

Page 8

32. I would consider leaving the random effect “experimenter” in the analysis whether or not including it creates statistically significant differences across models. There is always some variance explained by such random effects, even when not significant.

Response 32. OK will do! We revised the Methods and Analysis Plan to reflect this change.

Page 14

33. Have you decided to change the analyses after reading McElreath (2016)?

Response 33. Unfortunately, Logan has not had a chance to read the book yet because of all of the field work issues that came up in the past few months. She hopes to finish the book this summer and will update the preregistration accordingly before conducting any of the analyses in the analysis plan.

Page 15

34. I entirely understand the need for a day off in the experiments. However, does this mean that some individuals will experience a 1-day break during their experiments? Or between training and experiments? If so, this should be included in the analysis as a random factor, whenever possible.

Response 34. Yes, this does mean that some individuals experience a 1 day break in their experiments, and this break can also occur between training and the experiment. Great idea to add this as a random factor in the analyses! We have made the change to the preregistration throughout the Methods and the Analysis Plan sections.

Page 21

35. Here the training criterion is >1 item in at least 3 trials. In the detour task (page 22), the criterion is to correctly retrieve the food reward on the first attempt in 4 of 5 consecutive trials before receiving the test. In the go no-go task (page 24), the criterion is to retrieve food immediately in 8 out of the most recent 10 opportunities. Why use different criteria for the different experiments?

Response 35. For the detour and delay of gratification tests, we are trying to replicate the methods of MacLean et al. 2014 and Hillemann et al. 2014, respectively, as closely as possible, which is why we used their criterion for the grackles - to make our results more comparable with theirs. For the go no-go task, we used the criterion we use for the reversal learning experiments so that we were reducing the number of different criteria used across the various tests (when we are not replicating others' methods). We added notes to the protocol to make this clearer on pages 2 and 12 (highlighted in yellow; https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwcKOB_KcT7HE/edit?usp=sharing)

Page 22

36. Test: "subjects were allowed to retrieve the item on all trials regardless of the accuracy of their first attempt". With this method, the bird gets a reward both when it makes the correct choice and the incorrect choice. This has the potential to reduce learning drastically, since making a "wrong" choice only slightly delays access to the reward, but is still a rewarding choice. Is there a specific reason for letting the bird access the reward regardless of its performance in the test?

Response 36. We did this because we wanted to follow the methods used in MacLean et al. 2014 as closely as possible to make our data comparable with their large data set. In the warm-up phase, the bird learns where the food is and how to get it and as such is expected to get better at going to the side of the tube to retrieve the food over time (if it doesn't do this initially). However in the test trials, due to the transparent tube, individuals are expected to experience an impulse to go straight towards the reward and not around. The question this experiment addresses is not whether individuals can 'learn' to withhold the initial impulse to go straight towards the food, but rather it attempts to measure the initial tendency to do so, therefore individuals should always be rewarded, no matter which behaviour they show.

Page 24

37. If some of the birds experience other manipulations (or a significant time interval between the initial training and the test phase) whilst other birds are directly trained and tested, this has the potential to affect their learning abilities. If this is the case, I'd recommend giving a few training trials again to the birds which don't go straight from training to testing.

Response 37. Good catch. It turns out that all individuals in the flexibility manipulation condition and the flexibility control condition get touchscreen training in the same way and begin their touchscreen tests at the same time, so there is no difference in their training. We revised the protocol (highlighted in yellow on page 14 at https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwcKOB_KcT7HE/edit?usp=sharing) to reflect this. We also take your point that a bird may need a refresher from day to day, especially if they are just learning how to use the touchscreen. We revised the protocol (pages 15-16) to check the bird has retained the information from the previous day by running a few trials of the previous program before moving on to the next program.

Page 25

38. It is not clear when the bird receives the reward. Does it get rewarded immediately after pecking on the rewarded stimulus? Or at the end of the 10s presentation?

Response 38. Sorry for the confusion! The bird gets rewarded immediately after pecking the food key. We have now clarified this (highlighted in yellow on page 17 of the protocol at https://docs.google.com/document/d/1oEQ66yLrkMFr4UJTxfPBRAEXqoUuOgRwcKOB_KcT7HE/edit?usp=sharing).

39. Rewarding the bird at the end of the 10s presentation will also introduce a delay in getting the reward which can affect their abilities to associate the stimulus with the reward.

Response 39. Please see response 38.