



**Have you any concerns about statistical analyses in this paper?**

No

**Recommendation?**

Major revision is needed (please make suggestions in comments)

**Comments to the Author(s)**

Manuscript review comments are in the attached text file. See Appendix 1.

## Review form: Reviewer 2

**Is the manuscript scientifically sound in its present form?**

Yes

**Are the interpretations and conclusions justified by the results?**

Yes

**Is the language acceptable?**

Yes

**Is it clear how to access all supporting data?**

Yes

**Do you have any ethical concerns with this paper?**

No

**Have you any concerns about statistical analyses in this paper?**

No

**Recommendation?**

Accept with minor revision (please list in comments)

**Comments to the Author(s)**

Dear Author,

I find your work on the lack of innovation in novel tasks in grackles very interesting. It questions the assumption that behavioural flexibility and innovation can be used as interchangeable terms and helps to understand the multi-faceted nature of behavioural flexibility which opens up exciting future research avenues. There are, however, three points I would like you to address: A) Stone dropping experiment. This experiment is described elsewhere and should therefore be removed from Methods and Results. I do however think the results of this experiment can be a useful addition to the general discussion.

B) The methods section can be improved by adding more information about i) the experimental procedures and ii) the underlying rationale. I believe there were good reasons for all of this, but at first these are not always clear.

i. For example: how often did habituation (line 158, 200) and experimental trials (line 159, line 167, line 172-175, line 182, line 207) take place, i.e. how much time did pass between individual trials? This is important to know to understand if other reasons (e.g. lack of food motivation) could explain the results.

ii. For example: line 163/176: why would successful tool users have had 15 more trials, if the objective of this experiment was to confirm that they innovated stick tool use?

Line 203/226: why are trials for stone dropping and stick tool use 5 min long, but trials for horizontal string pulling only 3 minutes and for vertical string pulling 20 minutes?

C) The discussion needs to be broadened considerably. The new results are discussed in detail and also compared to the results from other species, but there is no discussion about where these differences might come from. It would also be interesting to discuss the difference in quality of behavioural actions between the experiments (e.g. interaction, but unsuccessful in experiment 2 and no interaction in experiment 3). The implications of your findings should be put in a more general context and linked back to examples from the literature, e.g. where behavioural flexibility and innovation has been used interchangeably.

Here are a few additional comments:

Line 50: spell out which aspects exactly, all three experiments here test if grackles invent new behaviours (line 63)

Line 59: add 'some' or make it more clear that grackles as a species are capable of showing behavioural flexibility in these tasks, but that there is large variation between individuals

Line 75: in comparison to? (other species that have been tested on this task?)

Line 77: see comment above. None of the other species would have experienced this task in the wild.

Line 80: define 'true tools'

Line 99: elaborate on the reason behind this prediction. The current phrasing suggests this prediction would hold true for all species

Line 119: were they visually isolated at all times?

Line 141: why did the stone dropping experiment not involve a similar habituation period to the apparatus as the other two experiments?

Line 197/198: immediately or within a certain time period?

Line 200: how long did a habituation trial last for?

Figure 2: add an introductory sentence to link this figure to the tool use experiment

Table 1: delete stone drop training trials column

Best wishes

## Decision letter (RSOS-160159)

01-Apr-2016

Dear Dr Logan:

Manuscript ID RSOS-160159 entitled "How far will a behaviourally flexible invasive bird go to innovate?" which you submitted to Royal Society Open Science, has been reviewed. The comments from reviewers are included at the bottom of this letter.

In view of the criticisms of the reviewers, the manuscript has been rejected in its current form. However, a new manuscript may be submitted which takes into consideration these comments.

Please note that resubmitting your manuscript does not guarantee eventual acceptance, and that your resubmission will be subject to peer review before a decision is made.

You will be unable to make your revisions on the originally submitted version of your manuscript. Instead, revise your manuscript and upload the files via your author centre.

Once you have revised your manuscript, go to <https://mc.manuscriptcentral.com/rsos> and login to your Author Center. Click on "Manuscripts with Decisions," and then click on "Create a Resubmission" located next to the manuscript number. Then, follow the steps for resubmitting your manuscript.

Your resubmitted manuscript should be submitted by 29-Sep-2016. If you are unable to submit by this date please contact the Editorial Office.

We look forward to receiving your resubmission.

Sincerely,  
Matthew Allinson,  
Editorial Coordinator, Royal Society Open Science

on behalf of

Kevin Padian, Royal Society Open Science

[openscience@royalsociety.org](mailto:openscience@royalsociety.org)

Associate Editor Comments to Author (Dr Matthew Allinson):

Associate Editor: 1

Comments to the Author:

If, upon resubmission, the manuscript under review with PeerJ has not yet been published, please submit a copy of that manuscript as part of the ESM to help the referees assess the methods.

Reviewers' Comments to Author:

Reviewer: 1

Comments to the Author(s)

Manuscript review comments are in the attached text file.

Reviewer: 2

Comments to the Author(s)

Dear Author,

I find your work on the lack of innovation in novel tasks in grackles very interesting. It questions the assumption that behavioural flexibility and innovation can be used as interchangeable terms and helps to understand the multi-faceted nature of behavioural flexibility which opens up exciting future research avenues. There are, however, three points I would like you to address:

A) Stone dropping experiment. This experiment is described elsewhere and should therefore be removed from Methods and Results. I do however think the results of this experiment can be a useful addition to the general discussion.

B) The methods section can be improved by adding more information about i) the experimental procedures and ii) the underlying rationale. I believe there were good reasons for all of this, but at first these are not always clear.

i. For example: how often did habituation (line 158, 200) and experimental trials (line 159, line 167, line 172-175, line 182, line 207) take place, i.e. how much time did pass between individual trials? This is important to know to understand if other reasons (e.g. lack of food motivation) could explain the results.

ii. For example: line 163/176: why would successful tool users have had 15 more trials, if the objective of this experiment was to confirm that they innovated stick tool use?

Line 203/226: why are trials for stone dropping and stick tool use 5 min long, but trials for horizontal string pulling only 3 minutes and for vertical string pulling 20 minutes?

C) The discussion needs to be broadened considerably. The new results are discussed in detail and also compared to the results from other species, but there is no discussion about where these differences might come from. It would also be interesting to discuss the difference in quality of behavioural actions between the experiments (e.g. interaction, but unsuccessful in experiment 2 and no interaction in experiment 3). The implications of your findings should be put in a more general context and linked back to examples from the literature, e.g. where behavioural flexibility and innovation has been used interchangeably.

Here are a few additional comments:

Line 50: spell out which aspects exactly, all three experiments here test if grackles invent new behaviours (line 63)

Line 59: add 'some' or make it more clear that grackles as a species are capable of showing behavioural flexibility in these tasks, but that there is large variation between individuals

Line 75: in comparison to? (other species that have been tested on this task?)

Line 77: see comment above. None of the other species would have experienced this task in the wild.

Line 80: define 'true tools'

Line 99: elaborate on the reason behind this prediction. The current phrasing suggests this prediction would hold true for all species

Line 119: were they visually isolated at all times?

Line 141: why did the stone dropping experiment not involve a similar habituation period to the apparatus as the other two experiments?

Line 197/198: immediately or within a certain time period?

Line 200: how long did a habituation trial last for?

Figure 2: add an introductory sentence to link this figure to the tool use experiment

Table 1: delete stone drop training trials column

Best wishes

## Author's Response to Decision Letter for (RSOS-160159)

See Appendix 2.

## RSOS-160247.R0 (Revision)

### Review form: Reviewer 1 (Berenika Mioduszezwska)

**Is the manuscript scientifically sound in its present form?**

Yes

**Are the interpretations and conclusions justified by the results?**

Yes

**Is the language acceptable?**

Yes

**Is it clear how to access all supporting data?**

It is clear how to access all supporting data.

**Do you have any ethical concerns with this paper?**

No

**Have you any concerns about statistical analyses in this paper?**

No

**Recommendation?**

Accept with minor revision (please list in comments)

**Comments to the Author(s)**

Dear Author,

I am very pleased to see how the previous feedback was implemented in your revised manuscript. However, I would still like to address some minor points (see my comments below).

**Background**

Lines 103 and 104: It seems that these tool use reports are from primates and dolphins ("objects are held in the hand or mouth", "hammers, sponges"), not birds as stated at the beginning of the sentence.

**Methods**

Line 164: "(...) because birds were allowed to peck holes..." would sound better.

Line 166: A copyright mark is missing when describing a commercially bought bird food. Also a clarification whether it was a seed mix, pressed granules etc would be helpful.

Line 170: When rolling in the experimental table, was the human experimenter also entering the aviary?

Line 171: If a human had to enter the experimental aviary, the wild grackles could still be stressed. It's different to be habituated to humans in a city vs in a small enclosure.

Line 174: Was the experimenter visible to the subject during testing?

Line 204: The tool use apparatus is also similar in design to the crevice platforms used in an ontogeny study on New Caledonian crows (Kenward et al, 2006; doi:10.1016/j.anbehav.2006.04.007). Which aspects of the methods/procedure from Auersperg et al were followed? The referred study used a different apparatus and different ghost control experimental conditions.

Line 244: The different types of strings were not introduced before. Re-phrasing this sentence into "(...) to two different types of strings (description below) by tying..." would make it easier to follow.

Line 269: Were the cups always red or was the colour important? The colour is only mentioned here and not before in Line 249 where the use of cups is introduced.

Line 288: A comma or a semicolon (without "grackles" capitalized) would be more appropriate to separate the two sentences inside the parentheses.

Line 293: Changing part of this sentence into "(...) made contact with the object or food reward" would make it easier to understand the procedure.

## Results

Lines 340 and 341: It is very strange that three individuals didn't pass the habituation/eat the bait. Did they differ in any way (more shy/neophobic etc) from the other subjects?

## Discussion

Line 353: The sex effect reference to Auersperg et al (2014) study is confusing as both female and male grackles interacted with the stick tool (data presented in Table 1).

Line 359: Goffin's cockatoos are likely to be feeding generalists, however currently we don't have literature evidence from the field to support this claim (see Auersperg et al, 2011, page 2, last paragraph of section "Subjects").

Line 363: Could grackles' failure to persist also stem from not being able to see the reward inside the slot apparatus (lacking large enough binocular overlap)? A slightly more detailed discussion of potential bill shape and vision effects on tool use capacity in grackles should include a study by Troscianko et al (2012, doi: 10.1038/ncomms2111) about New Caledonian crows' tool adapted bill morphology and binocular vision.

Line 364: For an interesting example of "awkward" stick tool use that could be added here, see the multi-access box study by Auersperg et al, 2011 (doi:10.1371/journal.pone.0020231) where all keas attempted to use stick tools but only one succeeded (ergonomically stick tool use is difficult for keas).

Line 403: One rook was also suggested to spontaneously drop a stone into the tube apparatus, however social influence cannot be excluded (Bird and Emery, 2009; doi: 10.1073/pnas.0901008106).

Line 424: I would re-phrase the last part of this sentence to something like: "species not known to use tools in the wild". There are anecdotal reports from the field of stick tool using keas (<http://www.stuff.co.nz/national/10417383/Sticky-beak-is-New-Zealands-tooled-up-kea>) and currently we don't know enough about the Goffin ecology and behaviour in the wild to exclude any potential tool use in this species.

Lines 436 and 437: Transition between these two paragraphs could be improved.

Line 466: "from grassy and recently plowed areas" sounds better than "from grassy areas and recently plowed areas".

Best,

Berenika Mioduszezewska

Max Planck Institute for Ornithology

## Review form: Reviewer 2

**Is the manuscript scientifically sound in its present form?**

Yes

**Are the interpretations and conclusions justified by the results?**

Yes

**Is the language acceptable?**

Yes

**Is it clear how to access all supporting data?**

Yes

**Do you have any ethical concerns with this paper?**

No

**Have you any concerns about statistical analyses in this paper?**

No

**Recommendation?**

Accept with minor revision (please list in comments)

**Comments to the Author(s)**

Dear Author,

I enjoyed reading your resubmitted manuscript. The re-write of introduction and discussion greatly enhanced the conciseness, ease to read and put the results in a more general context and the revisions resulted in a much clearer, more detailed and reproducible methods section. This is a useful contribution to the literature of behavioural flexibility and innovativeness.

A few very minor things (line numbers refer to the track changed version):

Line 65: I believe this is meant to say multi-faceted

Line 115-116: Clarify that this refers to the experiment described in this manuscript. The current phrasing (as it is part of the general introduction) could be misread as if this would be the case in the wild.

Line 300: The experiment numbers in the video clips do not longer correspond to the experiment numbers in the manuscript. If possible, remove the stone dropping clip from the video.

Best wishes.

**Decision letter (RSOS-160247)**

10-May-2016

Dear Dr Logan

On behalf of the Editor, I am pleased to inform you that your Manuscript RSOS-160247 entitled "How far will a behaviourally flexible invasive bird go to innovate?" has been accepted for publication in Royal Society Open Science subject to minor revision in accordance with the referee suggestions. Please find the referees' comments at the end of this email.

The reviewers and Subject Editor have recommended publication, but also suggest some minor revisions to your manuscript. Therefore, I invite you to respond to the comments and revise your manuscript.

- Ethics statement

If your study uses humans or animals please include details of the ethical approval received, including the name of the committee that granted approval. For human studies please also detail whether informed consent was obtained. For field studies on animals please include details of all permissions, licences and/or approvals granted to carry out the fieldwork.

- Data accessibility

It is a condition of publication that all supporting data are made available either as supplementary information or preferably in a suitable permanent repository. The data accessibility section should state where the article's supporting data can be accessed. This section should also include details, where possible of where to access other relevant research materials such as statistical tools, protocols, software etc can be accessed. If the data has been deposited in



an external repository this section should list the database, accession number and link to the DOI for all data from the article that has been made publicly available. Data sets that have been deposited in an external repository and have a DOI should also be appropriately cited in the manuscript and included in the reference list.

If you wish to submit your supporting data or code to Dryad (<http://datadryad.org/>), or modify your current submission to dryad, please use the following link:

<http://datadryad.org/submit?journalID=RSOS&manu=RSOS-160247>

- **Competing interests**

Please declare any financial or non-financial competing interests, or state that you have no competing interests.

- **Authors' contributions**

All submissions, other than those with a single author, must include an Authors' Contributions section which individually lists the specific contribution of each author. The list of Authors should meet all of the following criteria; 1) substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; 2) drafting the article or revising it critically for important intellectual content; and 3) final approval of the version to be published.

All contributors who do not meet all of these criteria should be included in the acknowledgements.

We suggest the following format:

AB carried out the molecular lab work, participated in data analysis, carried out sequence alignments, participated in the design of the study and drafted the manuscript; CD carried out the statistical analyses; EF collected field data; GH conceived of the study, designed the study, coordinated the study and helped draft the manuscript. All authors gave final approval for publication.

- **Acknowledgements**

Please acknowledge anyone who contributed to the study but did not meet the authorship criteria.

- **Funding statement**

Please list the source of funding for each author.

Because the schedule for publication is very tight, it is a condition of publication that you submit the revised version of your manuscript within 7 days (i.e. by the 19-May-2016). If you do not think you will be able to meet this date please let me know immediately.

To revise your manuscript, log into <https://mc.manuscriptcentral.com/rsos> and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions". Under "Actions," click on "Create a Revision." You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript and upload a new version through your Author Centre.

When submitting your revised manuscript, you will be able to respond to the comments made by the referees and upload a file "Response to Referees" in "Section 6 - File Upload". You can use this to document any changes you make to the original manuscript. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the referees.

When uploading your revised files please make sure that you have:

- 1) A text file of the manuscript (tex, txt, rtf, docx or doc), references, tables (including captions) and figure captions. Do not upload a PDF as your "Main Document".
- 2) A separate electronic file of each figure (EPS or print-quality PDF preferred (either format should be produced directly from original creation package), or original software format)
- 3) Included a 100 word media summary of your paper when requested at submission. Please ensure you have entered correct contact details (email, institution and telephone) in your user account
- 4) Included the raw data to support the claims made in your paper. You can either include your data as electronic supplementary material or upload to a repository and include the relevant doi within your manuscript
- 5) Included your supplementary files in a format you are happy with (no line numbers, vancouver referencing, track changes removed etc) as these files will NOT be edited in production

Once again, thank you for submitting your manuscript to Royal Society Open Science and I look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Best wishes

Andrew Dunn  
Senior Publishing Editor, Royal Society Open Science

on behalf of Kevin Padian  
Subject Editor, Royal Society Open Science  
openscience@royalsociety.org

Comments to Author:

Reviewer: 2

Comments to the Author(s)

Dear Author,

I enjoyed reading your resubmitted manuscript. The re-write of introduction and discussion greatly enhanced the conciseness, ease to read and put the results in a more general context and the revisions resulted in a much clearer, more detailed and reproducible methods section. This is a useful contribution to the literature of behavioural flexibility and innovativeness.

A few very minor things (line numbers refer to the track changed version):

Line 65: I believe this is meant to say multi-faceted

Line 115-116: Clarify that this refers to the experiment described in this manuscript. The current phrasing (as it is part of the general introduction) could be misread as if this would be the case in the wild.

Line 300: The experiment numbers in the video clips do not longer correspond to the experiment numbers in the manuscript. If possible, remove the stone dropping clip from the video.

Best wishes.

Reviewer: 1

Comments to the Author(s)

Dear Author,

I am very pleased to see how the previous feedback was implemented in your revised manuscript. However, I would still like to address some minor points (see my comments below).

#### Background

Lines 103 and 104: It seems that these tool use reports are from primates and dolphins ("objects are held in the hand or mouth", "hammers, sponges"), not birds as stated at the beginning of the sentence.

#### Methods

Line 164: "(...) because birds were allowed to peck holes..." would sound better.

Line 166: A copyright mark is missing when describing a commercially bought bird food. Also a clarification whether it was a seed mix, pressed granules etc would be helpful.

Line 170: When rolling in the experimental table, was the human experimenter also entering the aviary?

Line 171: If a human had to enter the experimental aviary, the wild grackles could still be stressed. It's different to be habituated to humans in a city vs in a small enclosure.

Line 174: Was the experimenter visible to the subject during testing?

Line 204: The tool use apparatus is also similar in design to the crevice platforms used in an ontogeny study on New Caledonian crows (Kenward et al, 2006; doi:10.1016/j.anbehav.2006.04.007). Which aspects of the methods/procedure from Auersperg et al were followed? The referred study used a different apparatus and different ghost control experimental conditions.

Line 244: The different types of strings were not introduced before. Re-phrasing this sentence into "(...) to two different types of strings (description below) by tying..." would make it easier to follow.

Line 269: Were the cups always red or was the colour important? The colour is only mentioned here and not before in Line 249 where the use of cups is introduced.

Line 288: A comma or a semicolon (without "grackles" capitalized) would be more appropriate to separate the two sentences inside the parentheses.

Line 293: Changing part of this sentence into "(...) made contact with the object or food reward" would make it easier to understand the procedure.

#### Results

Lines 340 and 341: It is very strange that three individuals didn't pass the habituation/eat the bait. Did they differ in any way (more shy/neophobic etc) from the other subjects?

#### Discussion

Line 353: The sex effect reference to Auersperg et al (2014) study is confusing as both female and male grackles interacted with the stick tool (data presented in Table 1).

Line 359: Goffin's cockatoos are likely to be feeding generalists, however currently we don't have literature evidence from the field to support this claim (see Auersperg et al, 2011, page 2, last paragraph of section "Subjects").

Line 363: Could grackles' failure to persist also stem from not being able to see the reward inside the slot apparatus (lacking large enough binocular overlap)? A slightly more detailed discussion of potential bill shape and vision effects on tool use capacity in grackles should include a study by Troscianko et al (2012, doi: 10.1038/ncomms2111) about New Caledonian crows' tool adapted bill morphology and binocular vision.

Line 364: For an interesting example of “awkward” stick tool use that could be added here, see the multi-access box study by Auersperg et al, 2011 (doi:10.1371/journal.pone.0020231) where all keas attempted to use stick tools but only one succeeded (ergonomically stick tool use is difficult for keas).

Line 403: One rook was also suggested to spontaneously drop a stone into the tube apparatus, however social influence cannot be excluded (Bird and Emery, 2009; doi: 10.1073/pnas.0901008106).

Line 424: I would re-phrase the last part of this sentence to something like: "species not known to use tools in the wild". There are anecdotal reports from the field of stick tool using keas (<http://www.stuff.co.nz/national/10417383/Sticky-beak-is-New-Zealands-tooled-up-kea>) and currently we don't know enough about the Goffin ecology and behaviour in the wild to exclude any potential tool use in this species.

Lines 436 and 437: Transition between these two paragraphs could be improved.

Line 466: "from grassy and recently plowed areas" sounds better than “from grassy areas and recently plowed areas”.

Best,  
Berenika Mioduszewska  
Max Planck Institute for Ornithology

## Author's Response to Decision Letter for (RSOS-160247)

See Appendix 3.

The submitted manuscript was interesting to read as it addressed the question of whether a behaviourally flexible bird species is also innovative in solving novel problem tasks. The posed question is of high value to the field of animal innovations and the style of the paper was easy to follow. A big advantage of the presented work is the use of multiple different tests, such as implementing both horizontal and vertical string pulling problem tasks. However, the manuscript would benefit greatly from enhancing the introduction and discussion of the results.

Below are my detailed comments on each section of the manuscript.

## Introduction

General: More information about grackles' ecology (such as where they live, what kind of environments they invaded, what is their social and feeding behaviour) would be very helpful in assessing the Author's predictions.

- Line 42: The given definition of behavioural flexibility is somehow confusing. It seems it contains parts of both the definitions of innovation proposed by Griffin and Guez 2014 (“innovations (...) ability to invent new behaviours, or use pre-existing ones in new contexts”) and behavioural flexibility (“Behavioural flexibility, including the ability to change behaviour in response to environmental feedback, and the ability to inhibit previously successful behaviour”; in: Griffin and Guez 2014). Following the behavioural flexibility definition Griffin and Guez state that it “determines the likelihood that new behaviours will emerge, and that pre-existing behaviours will arise in novel contexts” thus it determines the likelihood of innovations to occur. Additionally, on page 122 (under headline 2.1) Griffin and Guez state that: “several different processes should be critical to innovation, including novelty responses, exploration, asocial learning, and behavioural flexibility”. Therefore, it seems that these authors don't consider innovation and behavioural flexibility as synonymous but that behavioural flexibility is one of the many processes underlying innovation. However, in a recent review by Griffin (Innovativeness as an emergent property: a new alignment of comparative and experimental research on animal innovation. *Phil. Trans. R. Soc. B* 371: 20150544. <http://dx.doi.org/10.1098/rstb.2015.0544>) “behavioural flexibility” and “innovation” seem to be used as interchangeable terms... More discussion of the current trends in defining innovation vs behavioural flexibility would be useful.
- Line 58: High invasion rates might not necessarily mean that a species is highly behaviourally flexible, it depends on the novelty of the new environment (see: Griffin AS. 2016; chapter page 6: “Future directions in determining the ecological correlates of innovativeness”). A brief description of the environments to which grackles spread and in which aspects they are novel would be very useful.
- Line 85: A valuable point to add might be that keas which also turn stones (feeding generalist) are able to learn how to use tools (Auersperg, A. M. I., von Bayern, A. M. P., Gajdon, G. K., Huber, L., & Kacelnik, A. (2011). Flexibility in problem solving and tool use of kea and New

Caledonian crows in a multi access box paradigm. *PLoS One*, 6(6), e20231. doi:10.1371/journal.pone.0020231). More information about whether grackles are also feeding generalists would be useful to assess the ultimate ecological mechanisms at play.

## Methods

- Line 119: Could individual housing caused stress to the birds? More information about their social ecology is needed to appropriately evaluate the methods.
- Line 121: Was their exploration behaviour/motivation measured?
- Line 137: Was it preferred food or just any food from their daily diet? Information about the preferred food/reward would be useful to predict their motivation for participating in testing.
- Line 141: These seem like very short testing trials. Is there any data about grackles attention span, motivation levels, or whether they need to overcome initial neophobia when faced with an experimental situation (even after being habituated to the setup)?
- Line 156: Was there a reason that different rewards were used in training (peanuts) and tests (bread)? Wouldn't peanuts be a better reward also in testing?
- Line 181: I assume the grackles were habituated to a human presence prior to the social learning condition. It would be nice to see a one sentence description about the subjects' experience with humans.
- Line 198: Stylistics - "they were fed their dinner off of the inner table" was a bit confusing to read.
- Line 220: Was it an appropriate habituation for the strings? In previous experiments birds were often neophobic towards long hanging strings and various habituation protocols were used to overcome this response (see for example: Heinrich, B. and Bugnyar, T. (2005), Testing Problem Solving in Ravens: String-Pulling to Reach Food. *Ethology*, 111: 962–976. doi: 10.1111/j.1439-0310.2005.01133.x).

## Results

- Line 234: No birds dropped stones inside the tube but what was their behaviour, did they manipulate the stones, were they active around the apparatus or not active at all?
- Line 238: After the subjects pulled the stick out of the apparatus, did they continue to interact with it or lost interest?

## Discussion

General: This paper would benefit from adding a discussion about the grackles' ecology, neophobia, and motivation and how these might limit/confound their innovativeness.

Discussion of a potential age effect (see: Greenberg, R. (2012). The Role of Neophobia and Neophilia in the Development of Innovative Behaviour of Birds. *Animal Innovation*. doi:10.1093/acprof:oso/9780198526223.003.0008, page 184, "Uncurious adults") and possible factors influencing the grackles' performance in tests would be also welcome (e.g. for factors influencing performance in string pulling tasks see: Jacobs, I. F., & Osvath, M. (2015). The string-

pulling paradigm in comparative psychology. *Journal of Comparative Psychology*, 129(2), 89–120. doi:10.1037/a0038746). Maybe grackles are not as behaviourally flexible as was hypothesised in the previous study, or their flexibility is limited to certain contexts?

Also, it would be valuable to discuss how the fact that 2 out of 3 tasks used in this study required tool use from a non-tool using bird species could have influenced the results (see: Griffin 2016: “species that uses tools in the first place also has a greater knowledge base in the physical domain and broader skill set from which to invent new technical innovations through small behavioural variations”). Do these results provide a suggestion that tool use cannot be predicted by phylogeny only but rather by also looking at certain ecological circumstances in which tool use would be adaptive to evolve? An additional discussion of the ecology of grackles and why they don’t seem to need to use sticks for foraging would be welcome (see: Seed, A., & Byrne, R. (2010). Animal tool-use. *Current Biology*, 20(23), R1032–R1039. doi:10.1016/j.cub.2010.09.042, headline: “Tools and Evolution”). In general, the discussion of potential reasons why the grackles failed to use sticks and pull strings could be much improved.

- Table 1: Grackles needed large number of trials for training of dropping stones inside the tube apparatus. A discussion of how behavioural flexibility is/is not influenced by learning abilities would be informative in interpreting this result.
- Line 278: Why should a social learning condition work for grackles, do we know anything about their social learning/socio-ecology in the wild?
- Line 306: The references given here are quite old and coming from the same author (Lefebvre). A more modern and varied review of innovation and behavioural flexibility definitions would be useful to evaluate the Author’s proposal of not treating the innovation and flexibility as the same (see also a similar comment in the Introduction section).
- Line 309: As the reference about grackles’ foraging behaviour is a book from 1954 (not easily accessible), it would be nice to have some of the various feeding strategies listed here.

# Appendix A

## RESPONSE TO REVIEWERS AND EDITOR

Manuscript ID RSOS-160159

8 April 2016

Dear Dr. Allinson, Reviewer 1, and Reviewer 2,  
I cannot overstate how grateful I am for your detailed and helpful comments on the manuscript. Reviewer 2 had excellent restructuring/clarification feedback and Reviewer 1 went out of their way to suggest how to integrate more theory. I have revised the manuscript per your feedback and I provide responses to each comment below (marked by '>>>' and in italics). The line numbers listed below are from the tracked changes version (note that line numbers differ between this version and the clean version). I will renumber all of the references after the final revision. For now, the new references are referred to by the author's last name and year in the text and appear in alphabetical order at the end of the numbered references.

My PeerJ manuscript, on which some of the context in this paper rests, is now in press. It is not yet published so I placed the final revision in the ESM to allow reviewers to read it (titled: peerj-reviewing-9426-v2.pdf).

The manuscript has greatly benefitted from your feedback and I thank you for taking the time to give such useful comments! I look forward to hearing from you.

My best,  
Corina

### **REVIEWER 1**

The submitted manuscript was interesting to read as it addressed the question of whether a behaviourally flexible bird species is also innovative in solving novel problem tasks. The posed question is of high value to the field of animal innovations and the style of the paper was easy to follow. A big advantage of the presented work is the use of multiple different tests, such as implementing both horizontal and vertical string pulling problem tasks. However, the manuscript would benefit greatly from enhancing the introduction and discussion of the results.

>>> *Thank you!*

#### Introduction

General: More information about grackles' ecology (such as where they live, what kind of environments they invaded, what is their social and feeding behaviour) would be very helpful in assessing the Author's predictions.

>>> *I added more about their social and foraging behaviour and where they live (line 74-77, 160-166) to the introduction, and more about what environments they invaded in response to one of your comments below.*

- Line 42: The given definition of behavioural flexibility is somehow confusing. It seems it



contains parts of both the definitions of innovation proposed by Griffin and Guez 2014 (“innovations (...) ability to invent new behaviours, or use pre-existing ones in new contexts”) and behavioural flexibility (“Behavioural flexibility, including the ability to change behaviour in response to environmental feedback, and the ability to inhibit previously successful behaviour”; in: Griffin and Guez 2014). Following the behavioural flexibility definition Griffin and Guez state that it “determines the likelihood that new behaviours will emerge, and that pre-existing behaviours will arise in novel contexts” thus it determines the likelihood of innovations to occur. Additionally, on page 122 (under headline 2.1) Griffin and Guez state that: “several different processes should be critical to innovation, including novelty responses, exploration, asocial learning, and behavioural flexibility”. Therefore, it seems that these authors don’t consider innovation and behavioural flexibility as synonymous but that behavioural flexibility is one of the many processes underlying innovation. However, in a recent review by Griffin (Innovativeness as an emergent property: a new alignment of comparative and experimental research on animal innovation. Phil. Trans. R. Soc. B 371: 20150544. <http://dx.doi.org/10.1098/rstb.2015.0544>) “behavioural flexibility” and “innovation” seem to be used as interchangeable terms... More discussion of the current trends in defining innovation vs behavioural flexibility would be useful.

>>> *That is an excellent clarification, thank you. I think I unknowingly combined the definitions for innovation and behavioural flexibility (which, after re-reading, are clearly separate in Griffin and Guez 2014) because so often in the literature “innovation” is used interchangeably with “behavioural flexibility”. I have now updated the introduction to separate these terms as outlined in the Griffin & Guez paper and I contrasted this definition with previous definitions that treat the terms synonymously and placed this in the context of experimental results (lines 58-69).*

- Line 58: High invasion rates might not necessarily mean that a species is highly behaviourally flexible, it depends on the novelty of the new environment (see: Griffin AS. 2016; chapter page 6: “Future directions in determining the ecological correlates of innovativeness”). A brief description of the environments to which grackles spread and in which aspects they are novel would be very useful.

>>> *Yes, I have been thinking about this too. In introduced species, the new environment might be so similar to their previous environment that no flexibility is needed – they are already adapted. Native species that expand their range might do so because more of their suitable habitat is created or because they change their behaviour to access new niches. In the case of the great-tailed grackle, their suitable habitat (urban, grassy, and agricultural environments) has expanded as humans have. I have added a note about this and I discuss what aspects might have been novel at some point in their history (lines 79-109).*

- Line 85: A valuable point to add might be that keas which also turn stones (feeding generalist) are able to learn how to use tools (Auersperg, A. M. I., von Bayern, A. M. P., Gajdon, G. K., Huber, L., & Kacelnik, A. (2011). Flexibility in problem solving and tool use of kea and New Caledonian crows in a multi access box paradigm. PloS One, 6(6), e20231. doi:10.1371/journal.pone.0020231). More information about whether grackles are also feeding generalists would be useful to assess the ultimate ecological mechanisms

at play.

>>> *Great point, thank you! I added the kea example and that keas and grackles are generalist foragers (lines 125-159).*

## Methods

- Line 119: Could individual housing caused stress to the birds? More information about their social ecology is needed to appropriately evaluate the methods.

>>> *Initially, this was a concern of mine because I am used to working with long-term monogamous corvids. However, when I started observing the grackles in the field, I noticed that they did not have many affiliative behaviours (no allopreening, they rarely sit in proximity of each other) and they foraged in small groups that were loosely arranged. Later, when I brought them into the aviaries, I had planned on housing them together when they were not being tested, however they looked comfortable housed individually and I let them peck holes in the tarps at the back of their aviary so they could see each other (but not the testing section of the other's aviary). If I put them in the same aviary, there were extensive displacements and aggression such that I did not feel comfortable leaving them alone, so I kept them housed individually because it appeared less stressful. I added that this species is not monogamous and has weak social bonds, which is reported in the literature (lines 206-213).*

- Line 121: Was their exploration behaviour/motivation measured?

>>> *This text has been removed per comment A from Reviewer 2 below. Now the stone dropping experiment is only in the discussion because the data and results exist in a previous publication. Their exploration, risk aversion, neophobia, and motor diversity were examined (though not during stone dropping training) and related to behavioural flexibility in a manuscript that is in preparation for submission.*

- Line 137: Was it preferred food or just any food from their daily diet? Information about the preferred food/reward would be useful to predict their motivation for participating in testing.

>>> *This text has been removed per comment A from Reviewer 2 below. Now the stone dropping experiment is only in the discussion because the data and results exist in a previous publication. Good point though and it is clarified in the PeerJ manuscript where this data is published (the food was bread and/or peanuts so it was their preferred food that was only available during testing, not their maintenance diet).*

- Line 141: These seem like very short testing trials. Is there any data about grackles attention span, motivation levels, or whether they need to overcome initial neophobia when faced with an experimental situation (even after being habituated to the setup)?

>>> *This text has been removed per comment A from Reviewer 2 below. Now the stone dropping experiment is only in the discussion because the data and results exist in a previous*

publication. To answer your question, there is nothing published on their attention span or motivation levels because this species has never been tested in this way before. The spontaneous stone dropping trials were the same length as had been used on other species and they were short to avoid extinguishing the bird's interest in the apparatus entirely. They were habituated to the apparatus before the spontaneous stone dropping trials so their lack of innovation was likely not due to neophobia.

- Line 156: Was there a reason that different rewards were used in training (peanuts) and tests (bread)? Wouldn't peanuts be a better reward also in testing?

>>> *There was no particular reason that the habituation trial had a piece of peanut rather than a small piece of bread. Grackles liked both bread and peanuts and would work for both, with preferences sometimes appearing to change over time (I didn't measure this though). Bread was used for tool use testing because each piece could be made into a large size that blocked most of the opening, which would make its extraction with the stick easier. I removed bread and peanuts from the list of foods they were fed in an unrestricted way and noted that these two foods were used during testing (lines 213-216).*

- Line 181: I assume the grackles were habituated to a human presence prior to the social learning condition. It would be nice to see a one sentence description about the subjects' experience with humans.

>>> *The grackles were habituated to humans in general, and also their first 5 days in the aviary were spent habituating them to humans just outside their aviary door. I added the suggested description (lines 218-223).*

- Line 198: Stylistics - "they were fed their dinner off of the inner table" was a bit confusing to read.

>>> *I clarified the sentence (line 340).*

- Line 220: Was it an appropriate habituation for the strings? In previous experiments birds were often neophobic towards long hanging strings and various habituation protocols were used to overcome this response (see for example: Heinrich, B. and Bugnyar, T. (2005), Testing Problem Solving in Ravens: String-Pulling to Reach Food. *Ethology*, 111: 962-976. doi: 10.1111/j.1439-0310.2005.01133.x). Results

>>> *I used Heinrich & Bugnyar (2005) as the basis for my habituation protocols, which I noted on lines 310-311, and I have added more detail about how the grackles were habituated to strings (lines 369-371). The grackles behaved similarly to most of the ravens in Heinrich & Bugnyar (2005), even after both species had gone through habituation: They were neophobic of the strings hanging with objects attached (ravens "hesitatingly sidled up to the string and quickly jumped back after pecking it", pp.967-968). The difference between the species is that apparently the ravens eventually habituated to the set up (because all of them solved it) at some point after the experiments started (so some experimental time must have been devoted to habituation), whereas most grackles didn't. Perhaps grackles were less motivated and/or persistent in attempting to overcome their neophobia or, for those who were not neophobic, innovate the correct solution. I added this to the discussion and included the*

reference to Jacobs & Osvath 2015 (lines 515-523).

- Line 234: No birds dropped stones inside the tube but what was their behaviour, did they manipulate the stones, were they active around the apparatus or not active at all?

>>> *This text has been removed per comment A from Reviewer 2 below. Now the stone dropping experiment is only in the discussion because the data and results exist in a previous publication. They were active around the apparatus, walking around inspecting it from all angles, and 5 picked up and moved the stones or manipulated them.*

- Line 238: After the subjects pulled the stick out of the apparatus, did they continue to interact with it or lost interest?

>>> *I expanded the details on the range of interactions with the stick and noted that they lost interest in the stick after the interaction (lines 396-398).*

Discussion General: This paper would benefit from adding a discussion about the grackles' ecology, neophobia, and motivation and how these might limit/confound their innovativeness. Discussion of a potential age effect (see: Greenberg, R. (2012). The Role of Neophobia and Neophilia in the Development of Innovative Behaviour of Birds. *Animal Innovation*. doi:10.1093/acprof:oso/9780198526223.003.0008, page 184, "Uncurious adults") and possible factors influencing the grackles' performance in tests would be also welcome (e.g. for factors influencing performance in string pulling tasks see: Jacobs, I. F., & Osvath, M. (2015). The string-pulling paradigm in comparative psychology. *Journal of Comparative Psychology*, 129(2), 89–120. doi:10.1037/a0038746). Maybe grackles are not as behaviourally flexible as was hypothesised in the previous study, or their flexibility is limited to certain contexts?

>>> *I added to the discussion: a comparison between grackles and Goffin cockatoos regarding tool use and factors that might differ between grackles and other species that were successful (lines 454-488), age effects and neophilia with references to their ecology (lines 565-573). See my response to the comment starting with "Line 220" for a description of how I incorporated motivation and persistence (including referencing Jacobs & Osvath 2015) into the discussion. I clarified that the innovation results do not call into question previous behavioural flexibility results because these two traits are independent (lines 539-545).*

Also, it would be valuable to discuss how the fact that 2 out of 3 tasks used in this study required tool use from a non-tool using bird species could have influenced the results (see: Griffin 2016: "species that uses tools in the first place also has a greater knowledge base in the physical domain and broader skill set from which to invent new technical innovations through small behavioural variations"). Do these results provide a suggestion that tool use cannot be predicted by phylogeny only but rather by also looking at certain ecological circumstances in which tool use would be adaptive to evolve? An additional discussion of the ecology of grackles and why they don't seem to need to use sticks for foraging would be welcome (see: Seed, A., & Byrne, R.

(2010). Animal tool-use. *Current Biology*, 20(23), R1032–R1039.

doi:10.1016/j.cub.2010.09.042, headline: “Tools and Evolution”). In general, the discussion of potential reasons why the grackles failed to use sticks and pull strings could be much improved.

>>> *I added a discussion about why grackles might not need to invent tool use in the wild and what an investigation of a non-tool user might tell us about innovation (lines 546-564).*

- Table 1: Grackles needed large number of trials for training of dropping stones inside the tube apparatus. A discussion of how behavioural flexibility is/is not influenced by learning abilities would be informative in interpreting this result.

>>> *I added a discussion about how learning to drop stones should not influence the flexibility experiments, which were conducted in a different context (water tubes) and were based on choices between novel objects in changing circumstances (lines 539-545).*

- Line 278: Why should a social learning condition work for grackles, do we know anything about their social learning/socio-ecology in the wild?

>>> *I added to the introduction that grackles may use social information because they forage in flocks and, when an innovation occurs, it appears to spread to many individuals (lines 160-166). I figure it will be better for readers to have this context sooner rather than at the end in the discussion.*

- Line 306: The references given here are quite old and coming from the same author (Lefebvre). A more modern and varied review of innovation and behavioural flexibility definitions would be useful to evaluate the Author’s proposal of not treating the innovation and flexibility as the same (see also a similar comment in the Introduction section).

>>> *I added that these results support the revised framework by Griffin & Guez (2014), making it consistent with the revised introduction (lines 588-590).*

- Line 309: As the reference about grackles’ foraging behaviour is a book from 1954 (not easily accessible), it would be nice to have some of the various feeding strategies listed here.

>>> *I’m very happy to add details here; I found the read fascinating! Lines 592-595.*

## **REVIEWER 2**

Dear Author,

I find your work on the lack of innovation in novel tasks in grackles very interesting. It questions the assumption that behavioural flexibility and innovation can be used as interchangeable terms and helps to understand the multi-faceted nature of behavioural flexibility which opens up exciting future research avenues. There are, however, three points I would like you to address:

>>> *Thank you!*

A) Stone dropping experiment. This experiment is described elsewhere and should therefore be removed from Methods and Results. I do however think the results of this experiment can be a useful addition to the general discussion.

>>> *I had debated about this so I am glad you have a clear opinion on what to do with the previous findings. I removed stone dropping from the introduction, methods, and results, and I rephrased it in the discussion and abstract.*

B) The methods section can be improved by adding more information about i) the experimental procedures and ii) the underlying rationale. I believe there were good reasons for all of this, but at first these are not always clear.

>>> *I addressed your two comments on this topic below and also implemented an additional clarification:*

*Horizontal string pulling: if a bird did not come to the table in 3 min, a break was given (lasting from 5 min to 3 days) and the next trial was preceded by a habituation trial. Also, habituation was restarted if the bird appeared to be having a neophobic reaction (lines 348-352).*

i. For example: how often did habituation (line 158, 200) and experimental trials (line 159, line 167, line 172-175, line 182, line 207) take place, i.e. how much time did pass between individual trials? This is important to know to understand if other reasons (e.g. lack of food motivation) could explain the results.

>>> *I explained that the vertical string habituation trials ranged between 3-8 for those birds that never passed (Refresco=3 sessions in 1 day, Batido=3 sessions in 2 days, Margarita=5 sessions in 9 days, Michelada=8 sessions in 9 days), and between 1-2 trials for birds that passed and moved onto the experimental trials (Cerveza and Horchata=1 session in 1 day, Tequila=2 sessions in 2 days) (lines 374-377). I added that the horizontal string habituation trials ranged between 2-6 sessions in 2-5 days for birds that passed (Batido=2 sessions in 3 days, Horchata=3 sessions in 3 days, Refresco=2 sessions in 2 days, Michelada 5 sessions in 2 days, Tequila=4 sessions in 3 days, Margarita=3 sessions in 2 days, Cerveza=6 sessions in 5 days). Jugo didn't pass habituation, but he only got 1 session in 1 day because I ran out of time and needed to release them (lines 336-338). I explained that tool use experimental trials began immediately after passing habituation (line 278), that tool use experimental trials were given one immediately after the other (lines 288-289, 294, 310), and that horizontal string pulling trials could occur one immediately after the other or on consecutive days (lines 352).*

ii. For example: line 163/176: why would successful tool users have had 15 more trials, if the objective of this experiment was to confirm that they innovated stick tool use?

>>> *I added that 15 more trials were conducted to determine whether the bird's success was due to chance or whether it could consistently solve the task (lines 282-284).*

Line 203/226: why are trials for stone dropping and stick tool use 5 min long, but trials for horizontal string pulling only 3 minutes and for vertical string pulling 20 minutes?

>>> *I added that I based the methods for each experiment on previous studies that had been conducted on other species to make the grackle results comparable (lines 227-228, 314-315). The methods for stone dropping were removed per your comment A.*

C) The discussion needs to be broadened considerably. The new results are discussed in detail and also compared to the results from other species, but there is no discussion about where these differences might come from. It would also be interesting to discuss the difference in quality of behavioural actions between the experiments (e.g. interaction, but unsuccessful in experiment 2 and no interaction in experiment 3). The implications of your findings should be put in a more general context and linked back to examples from the literature, e.g. where behavioural flexibility and innovation has been used interchangeably.

>>> *The discussion has been greatly expanded per comments from Reviewer 1, which are in line with your comment here about where these differences might come from (see above responses starting at "Discussion General"). I added a discussion about the quality of the actions in the tool use (lines 481-488) and string pulling experiments (lines 515-523). I added a section to the introduction that discusses how empirical findings contradict the idea that behavioural flexibility and innovation are the same trait (lines 58-69), and I linked the discussion back to the general context that is now more fleshed out in the introduction (lines 588-591).*

Here are a few additional comments:

Line 50: spell out which aspects exactly, all three experiments here test if grackles invent new behaviours (line 63)

>>> *This has now been removed after clarifying theory in the introduction per Reviewer 1's comments.*

Line 59: add 'some' or make it more clear that grackles as a species are capable of showing behavioural flexibility in these tasks, but that there is large variation between individuals

>>> *Yes, good point. I changed this (line 109).*

Line 75: in comparison to? (other species that have been tested on this task?)

>>> *This has now been removed due to the revision from your comment A.*

Line 77: see comment above. None of the other species would have experienced this task in the wild.

>>> *This has now been removed due to the revision from your comment A.*

Line 80: define 'true tools'

>>> *I added a definition on lines 116-117.*

Line 99: elaborate on the reason behind this prediction. The current phrasing suggests this prediction would hold true for all species

>>> *The species that have exhibited string pulling behaviour are nest building species (including bumblebees), therefore perhaps this experience gave them the necessary skills to transfer to a new context. I expanded on this (lines 177-185).*

Line 119: were they visually isolated at all times?

>>> *No, they often pecked holes in the tarps that covered the walls for visual isolation during testing. I allowed holes in the tarps in areas where the adjacent bird would not have visual access to the testing part of its neighbours cage. Therefore, they could often see each other at the back of the aviaries, but not at the front where testing occurred. I clarified this (lines 210-213).*

Line 141: why did the stone dropping experiment not involve a similar habituation period to the apparatus as the other two experiments?

>>> *They were habituated to the apparatus before the spontaneous stone dropping experiment; I had just forgotten to mention this. However, this has now been removed in response to your comment A.*

Line 197/198: immediately or within a certain time period?

>>> *I added that the habituation period started if the bird did not eat the peanut after a few minutes (line 339).*

Line 200: how long did a habituation trial last for?

>>> *A bird was considered habituated if it ate the peanut within the first minute of the habituation trial. I added this to line 342.*

Figure 2: add an introductory sentence to link this figure to the tool use experiment

>>> *I added "To determine whether grackles innovate string pulling behaviour, two string pulling experiments were conducted: Horizontal (A, B, C) and vertical (D)". I also added a similar first sentence to Figure 1 to make them consistent.*

Table 1: delete stone drop training trials column

>>> *I deleted it.*



# Appendix A

## RESPONSE TO REVIEWERS AND EDITOR

Manuscript ID RSOS-160159

11 May 2016

Dear Andrew Dunn, Kevin Padian, Berenika Mioduszezewska, and Reviewer 2, I am so glad you think my revision was effective! I greatly appreciate your further comments here, which really helped polish the paper. I have revised the manuscript per your feedback and I provide responses to each comment below (marked by '>>>' and in italics). The line numbers refer to the tracked changes version (which are different from the line numbers in the clean version). Additionally, I added a more nuanced discussion of how invasion success is predicted to relate with behavioral flexibility in the Introduction after finding a couple new papers on this topic (L62-66).

Thank you again!

My best,  
Corina

### **REVIEWER 1**

I am very pleased to see how the previous feedback was implemented in your revised manuscript. However, I would still like to address some minor points (see my comments below).

*>>> Thank you very much! I'm happy to address your further comments and I appreciate the time you have taken to help me improve the manuscript!*

#### Background

Lines 103 and 104: It seems that these tool use reports are from primates and dolphins ("objects are held in the hand or mouth", "hammers, sponges"), not birds as stated at the beginning of the sentence.

*>>> Good point. I found text in the reference that specifically refers to birds holding a tool in the beak or foot and I revised my text (L78-79).*

#### Methods

Line 164: "(...) because birds were allowed to peck holes..." would sound better.

*>>> Thanks! I made the change (L140).*

Line 166: A copyright mark is missing when describing a commercially bought bird food. Also a clarification whether it was a seed mix, pressed granules etc would be helpful.

*>>> I added a registered trademark after Mazuri and described that they were pressed granules (L142).*

Line 170: When rolling in the experimental table, was the human experimenter also entering the

aviary?

>>> *Yes, I noted this (L146).*

Line 171: If a human had to enter the experimental aviary, the wild grackles could still be stressed. It's different to be habituated to humans in a city vs in a small enclosure.

>>> *True. Grackles displayed a range of disturbance upon experimenters entering the aviary, however all grackles calmed down almost immediately upon the experimenter leaving the aviary even though the experimenter was still in full view just outside the wire mesh door. I clarified this on L147-149.*

Line 174: Was the experimenter visible to the subject during testing?

>>> *Yes, this is noted on L150-151.*

Line 204: The tool use apparatus is also similar in design to the crevice platforms used in an ontogeny study on New Caledonian crows (Kenward et al, 2006; doi:10.1016/j.anbehav.2006.04.007). Which aspects of the methods/procedure from Auersperg et al were followed? The referred study used a different apparatus and different ghost control experimental conditions.

>>> *Thanks for pointing out Kenward et al.'s crevice apparatus! I added the reference (L160-161). I clarified that my trial durations, conditions, and passing criterion were modeled after Auersperg et al. (L160-161).*

Line 244: The different types of strings were not introduced before. Re-phrasing this sentence into "(...) to two different types of strings (description below) by tying..." would make it easier to follow.

>>> *Good catch! I changed the text as suggested (L204).*

Line 269: Were the cups always red or was the colour important? The colour is only mentioned here and not before in Line 249 where the use of cups is introduced.

>>> *The cups were always red so I removed the reference to their color on L230.*

Line 288: A comma or a semicolon (without "grackles" capitalized) would be more appropriate to separate the two sentences inside the parentheses.

>>> *I changed it to a semicolon (L250).*

Line 293: Changing part of this sentence into "(...) made contact with the object or food reward" would make it easier to understand the procedure.

>>> *Good point. I made the change (L255).*

## Results

Lines 340 and 341: It is very strange that three individuals didn't pass the habituation/eat the bait. Did they differ in any way (more shy/neophobic etc) from the other subjects?

>>> *Jugo, Batido, and Refresco did not differ consistently from Horchata in exploration behavior, and Horchata was the most neophobic of the four of them (these results are in a paper that is in review at PeerJ). The string pulling tests were done near the end of their time in the aviary, which was near the beginning of the breeding season and the males had begun to vocalize. Anecdotally, it seemed like they were still willing to participate in experiments that they were already familiar with, but were less inclined to learn about new experiments as the breeding season approached. I added a note about this (L302-305).*

## Discussion

Line 353: The sex effect reference to Auersperg et al (2014) study is confusing as both female and male grackles interacted with the stick tool (data presented in Table 1).

>>> *I eliminated the description of the Goffin cockatoo sex differences to clarify how the cockatoos are similar to the grackles (L310-311).*

Line 359: Goffin's cockatoos are likely to be feeding generalists, however currently we don't have literature evidence from the field to support this claim (see Auersperg et al, 2011, page 2, last paragraph of section "Subjects").

>>> *Thanks for the reference and for clarifying what is and is not known about the Goffin's cockatoos. I have the Auersperg et al. 2014 reference in there because I think I had seen her side note in the Subjects section (L315).*

Line 363: Could grackles' failure to persist also stem from not being able to see the reward inside the slot apparatus (lacking large enough binocular overlap)? A slightly more detailed discussion of potential bill shape and vision effects on tool use capacity in grackles should include a study by Troscianko et al (2012, doi: 10.1038/ncomms2111) about New Caledonian crows' tool adapted bill morphology and binocular vision.

>>> *Good idea to give a bit more detail here and to cite Trocianko et al. (2012) (L320-326).*

Line 364: For an interesting example of "awkward" stick tool use that could be added here, see the multi-access box study by Auersperg et al, 2011 (doi:10.1371/journal.pone.0020231) where all keas attempted to use stick tools but only one succeeded (ergonomically stick tool use is difficult for keas).

>>> *Interesting, I hadn't realized that paper hypothesized about bill morphology as well. I added a note about this on L325-326.*

Line 403: One rook was also suggested to spontaneously drop a stone into the tube apparatus, however social influence cannot be excluded (Bird and Emery, 2009; doi:

10.1073/pnas.0901008106).

>>> *I didn't realize this; thanks for pointing it out. I amended the sentence (L367-369).*

Line 424: I would re-phrase the last part of this sentence to something like: "species not known to use tools in the wild". There are anecdotal reports from the field of stick tool using keas (<http://www.stuff.co.nz/national/10417383/Sticky-beak-is-New-Zealands-tooled-up-kea>) and currently we don't know enough about the Goffin ecology and behaviour in the wild to exclude any potential tool use in this species.

>>> *I made the suggested change (L393).*

Lines 436 and 437: Transition between these two paragraphs could be improved.

>>> *I added a concluding sentence to smooth the transition (L406-407).*

Line 466: "from grassy and recently plowed areas" sounds better than "from grassy areas and recently plowed areas".

>>> *I made the suggested change (L430).*

## **REVIEWER 2**

I enjoyed reading your resubmitted manuscript. The re-write of introduction and discussion greatly enhanced the conciseness, ease to read and put the results in a more general context and the revisions resulted in a much clearer, more detailed and reproducible methods section. This is a useful contribution to the literature of behavioural flexibility and innovativeness.

>>> *Thank you so much!*

A few very minor things (line numbers refer to the track changed version):

Line 65: I believe this is meant to say multi-faceted

>>> *Good catch! I made the change (L51).*

Line 115-116: Clarify that this refers to the experiment described in this manuscript. The current phrasing (as it is part of the general introduction) could be misread as if this would be the case in the wild.

>>> *I made the clarification (L92).*

Line 300: The experiment numbers in the video clips do not longer correspond to the experiment numbers in the manuscript. If possible, remove the stone dropping clip from the video.

>>> *I revised the video to match the revised paper and I will re-upload it when I have more details about the citation to include at the end.*