Collated Peer Reviews for manuscript "*Are blue tits more likely to use social information from live demonstrators compared to video playback?*"

*Note that this collation includes the initial review of original manuscript (**MS**), plus the final assessment of revised manuscript (**RS**). Process is here in reversed chronological order, i.e. final assessments are first and initial peer reviews further down.*

FINAL ASSESSMENT OF REVISED MANUSCRIPT

Final assessment by Peer 4521 [see initial review]

Publishable

The authors have well answered all the questions and comments from me and other reviewers. And they have revised the manuscript according to these suggestions.

The Figure 1 is very informative. There has recently been discussion that articles of animal behaviour have less images than before of the observed behaviour and/or study set-up (Klein & Seeley 2015 Animal Behaviour).

The text of the manuscript is now much better. I have only two small comments:

lines 178, 197, 225-226, 227: replace x with × (multiplication sign)

line 320: cardboard

Qualitative scores on RS value:	
(scale: 1 [low] - 5 [high])	
Breadth:	4
Impact:	4
Originality:	4
Technical scores on RS soundness:	
(scale: 1 [poor] - 5 [excellent])	
Methods:	4
Data:	4
Inference:	4
Writing:	5

Final assessment by Peer 4519 [see initial review]

Publishable

This is an interesting manuscript that is carefully prepared and well presented. It has further improved by the revision. The results are not outstandingly exciting, but worth to be read by the community of people using video demonstrations in behavioural assays.

Qualitative scores on RS value:	
(scale: 1 [low] - 5 [high])	
Breadth:	2
Impact:	2
Originality:	4
Technical scores on RS soundness:	
(scale: 1 [poor] - 5 [excellent])	
Methods:	4
Data:	4
Inference:	4
Writing:	5

Final assessment by Peer 4518 [see initial review]

peer did not provide final assessment

INITIAL PEER REVIEWS ON ORIGINAL MANUSCRIPT

Note that in Peerage of Science, peer reviews themselves are also peer-reviewed when more than one reviewer is engaged: each initial review entry begins with peer-review-of-peer-review evaluation showing how accurate and justified other Peers consider that text to be on merits, critique and discussion.

Review by Peer 4521 [see final assessment]

PEER-REVIEW-OF-PEER-REVIEW

Evaluation of this review (average over Peers): PEQ: 4.7 Weight relative to best review: 1.00	**Scoring by section** *(scale: 1 [poor] - 5 [excellent])* *Merits: 5.0* *Critique: 5.0* *Discussion: 4.0*
Evaluation of this review by Peer 4518: Very good comments with statement the authors should present data on light conditions. - I think that the reviewer has justified comments for this paper.	*Merits: 5* *Critique: 5* *Discussion: 5*
Evaluation of this review by Peer 4519: Both points in the critique are valuable. However, I think the flickering frequency of the monitor would be more relevant if there was an actual differences between live demonstrators and videos. -	*Merits: 5* *Critique: 5* *Discussion: 3*

Introduction

Revision Recommendations on MS:	
Question:	Accept
Data:	Accept
Methods:	Minor Revision
Inference:	Accept
Writing:	Accept

The authors of "Are blue tits more likely to use social information from 1 live demonstrators compared to video playback?" studied the responses of blue tits when they could observe the behaviour of live conspecifics or video playback of a similar situation. The results were not as straight-forward as the researchers had probably hoped for. Birds always preferred the square symbol (and not cross or white symbols) despite what the demonstrator bird had used. Blue tits behaved similarly after observing live conspecific or video playback, so for this species using video playbacks in future experiments could work.

Merits

The study points out a critical problem if animals behave similarly after watching live conspecific or a video playback of a similar situation. Not all animals "understand" videos, so it's important to study their responses in both situations.

The study set-up seems well planned and conducted.

The language of the manuscript is good.

Critique

Blue tits can distinguish light-dark cycles up to 145 Hz which means that a light source with lower frequency is not seen as constant but "separate pictures" (Boström et al. 2016). You mention this possible problem with flicker-fusion frequency on lines 66-67. Therefore, it would be very important to tell more about the lighting conditions and video playback details in this experiment. Did the birds see the videos as a constant film or as a chain of separate pictures?

The results of this experiment leave a question that was the choice of used symbols the best possible. Maybe square and cross were too different. Would the results have been "better" if the symbols had been square and round, or square and triangle? Then the area of the symbol could have been the same. I'm also not totally sure why the authors wanted to include also a white symbol. Would it have been easier to analyse the behaviour of the birds if they had to choose only between two different symbols?

Discussion

As I mentioned already earlier, this is a very necessary study to compare the responses of blue tits to either live conspecific or video playback. The authors compare well their results to same kind of studies done with other species.

References

Boström et al. 2016. Ultra-Rapid Vision in Birds. PLoS ONE 11(3): e0151099

Additional Comments

lines 16-18: Is it worth mentioning this previous experiment already in abstract? Maybe Introduction should be a better place.

lines 23-24: Did the tray to control birds observed, had similar symbols as the tray the demonstrators had?

lines 24-26: The 'against' in the beginning of the sentence makes it difficult to understand this sentence. Tell first what is your hypothesis and after it that birds did not behave according to it.

lines 75, 92, 344, 351: scientific names of California scrub-jay, great tit, zebra finch, Burmese red junglefowl

line 160: Did you record the sex of the studied birds? If yes, did it have any effect on the behaviour?

line 241: Give some explanation why it was ok to combine live and video education?

line 309: There could be more discussion about how stressful the study situation was for the observer birds.

Fig. 2 a): The colour dark grey is too close to the colour black. Use lighter grey or some other colour combination for easier observation.

Review by Peer 4519 [see final assessment]

PEER-REVIEW-OF-PEER-REVIEW

Evaluation of this review (average over Peers): PEQ: 3.8 Weight relative to best review: 0.82	**Scoring by section** *(scale: 1 [poor] - 5 [excellent])* *Merits: 4.5* *Critique: 4.0* *Discussion: 3.0*
Evaluation of this review by Peer 4518: I agree with the reviewer that it could be that the absence of copying other blue tit individuals could be that they could draw the conclusion that they avoid areas where other individuals have been foraging, because these sites could have less density of prey after other individuals have exploited these sites. However, if it is a really good patch, with lots of food items, then it should pay to copy the behavior and forage on these sites. - I think that the reviewer has a good point that the experiment is what makes this paper interesting to read, but I think it could be to early to say if it could be of general interest to a broader audience of readers, it could give inspiration to others to investigate this question in many other species of birds, which I believe what be worthful.	*Merits: 5* *Critique: 5* *Discussion: 3*
Evaluation of this review by Peer 4521: The reviewer raised some points that I hadn't thought of. I don't think blue tits are that solitary foragers. Maybe to an extent during breeding but otherwise they spend time in flocks of con- and interspecifics. Therefore, they should find it useful to use social information of places with plenty of food. - I agree with the merits but not all the critique.	*Merits: 4* *Critique: 3* *Discussion: 3*

Introduction

Revision Recommendations on MS:	
Question:	Accept
Data:	Accept
Methods:	Accept
Inference:	Accept
Writing:	Minor Revision

The manuscript aims to answer the questions whether blue tits respond differently to video playbacks as compared to demonstrations by live conspecifics. The study is embedded in a context of social information use in a foraging. It is not entirely clear to me why solitary foragers of dispersed prey like blue tits should use social information, since places where conspecifics have been hunting successfully are likely to be less rewarding than places that have not been searched. The alternative expectation that blue tits avoid place for demonstrators have been successful seems equally valid to me. Unfortunately, the two effects may cancel each other resulting in no use of social information in foraging. The negative result reported in the manuscript has been found

by the authors in the same study system using video playbacks [1]. The current manuscript now shows that the negative outcome does not depend on whether video playbacks or live demonstrators are used. With 40 individuals being tested (20 per demonstration treatment), the sample size is decent, but also not very impressive, making lack of statistical power a potential concern. Hence the study remains a bit inconclusive in the end.

Merits

The manuscript poses the relevant question how experimental demonstrations should be performed when studying social interactions and learning. The manuscript is very clearly written and easy to read. The analyses have been competently done and background and methods are clearly described.

Critique

There is only one main problem with the study and this is the lack of a positive control. Since blue tits do not seem to respond neither to live demonstrators nor to video playbacks, we do not know whether the two approaches differ significantly from each other. It would have been nice to see the setup being tested in a context where there is a response when exposed to live demonstrations. On the one hand, it is interesting to see that blue tits do not seem to respond to demonstrations and that the conclusions of the original study [1] is robust. On the other hand, the difference between live demonstrations and video playbacks is likely to be dependent on context, making it difficult to draw general conclusions based on a single study.

I acknowledge that the authors report a stronger preference for squares after exposure to demonstrators, but this reads like a test that arose after data inspection. Unfortunately, the unequal strength of the stimuli has made it difficult to identify a clear overall effect of social information on foraging decisions.

Discussion

Overall, I think this is a very readable and concise manuscript that poses a relevant question. However, the question is a little too big for this study, so that we end up with a result that is not finally conclusive. The main achievement of the manuscript is thus that it raises a relevant question. The study is probably also relevant for the researchers studying blue tits (of which there are quite a few), since I believe the results are informative for designing future experiments with *Parus* tits. Beyond this group, and certainly beyond passerine birds, I think the manuscript is of rather limited interest. After all, I enjoyed the read and the insights into experimental work with blue tits.

References

[1] Hämäläinen, L., Rowland, H. M., Mappes, J., & Thorogood, R. (2017). Can video playback provide 420 social information for foraging blue tits? *PeerJ*, 5, e3062.

Additional Comments

It would be nice to read a statement on whether the size of demonstrators in video playback was calibrated to a realistic size of a live bird.

L49: Suggest "As the number of …"

L55: remove "or"

L77: "observers" should probably be "demonstrators"

L125: Suggest to move "in their home cages" to the end of the sentence.

L191: Food restriction was applied for how long?

L215: "was the same"

L217: remove "therefore"

L261: Remove "random". By chance does not need to be qualified by randomness.

L282: Suggest to start the discussion with a brief repetition of the research questions. Something like "We here test if …"

Figure 2 is lacking proper axis labels.

Review by Peer 4518 [see final assessment]

PEER-REVIEW-OF-PEER-REVIEW

Evaluation of this review (average over Peers): PEQ: 4.0 Weight relative to best review: NA	**Scoring by section** *(scale: 1 [poor] - 5 [excellent])* *Merits: 4.5* *Critique: 3.5* *Discussion: 4.0*
Evaluation of this review by Peer 4519: It is not actually clear if the critique refers to the sample size (in terms of individuals being tested) or to the number of occasions to learn form (hence the number of demonstration events). I agree, both aspects may be relevant in the present manuscript. -	*Merits: 5* *Critique: 3* *Discussion: 4*
Evaluation of this review by Peer 4521: The reviewer raises up valid merits and critique. Especially the point mentioned in critique is something that should be discussed in the paper. - I agree with this review.	*Merits: 4* *Critique: 4* *Discussion: 4*

Revision Recommendations on MS:
Question: Minor Revision

Data:	Accept
Methods:	Accept
Inference:	Accept
Writing:	Accept

Introduction

This paper on blue tits is about using playback experiments to study if birds observing other birds, the so-called demonstrators, and whether they can use that information to easier find food in a lab. experiment. They did not find any evidence that birds copied a demonstrator´s choice, regardless if social information was presented by observers being able to see the demonstrator´s directly, or where observers were presented with a video playback of the demonstrators foraging behavior. However, they showed that social information although had an influence on blue tit's foraging choices, in that socially educated birds, seemed to form a stronger preference for a square symbol rather than cross o plain white symbols.

Merits

The merits of this manuscript is the experimental comparison using demponstrators and then filming each demonstrator, and comparing this with a live demonstrator and a control. It also is stimulating more research in this area, as there still need to be more studies to disentangle effects of video playback on foraging behavior comparing controls and birds having seen a demonstrators foraging behavior.

Critique

The biggest critque I have of on this manuscript, is the sample size in the experiment. My first impression when reading the paper, was that it might be too few individuals to match a live sitution, where an observer could most probably follow and learn from many demonstrators in a population. Thus, there may be a tipping point in the copying of a behavior, such that when a bird, like the blue tit, have been observing a demonstrator´s foraging behavior a lot of times, then it will copy that behavior to a much larger extent, than it was found in this study.

Discussion

The importance of this paper is that it is shown in the study that blue tits reacted after seeing the video playback, but not in the predicted manner, and they do an interesting discussion about this phenomenon. As the authors discuss in the last part of the discussion, "individuals are also likely to vary in their tendency to use social information, and we might have needed a bigger sample size to detect social learning". I think that it could be quite a large variation in how individuals respond to social information, and one way to improve research in this area, is therefore to have a larger sample size, for example, that observers could be from different locations in the country where the study is performed.

References

There are many good references in the manuscript, and from a relatively long period of time of the different references.

Additional Comments

I have no furhter comments on the text.