THE ROYAL SOCIETY PUBLISHING

PROCEEDINGS B

High temperatures drive offspring mortality in a cooperatively breeding bird

Amanda R. Bourne, Susan J. Cunningham, Claire N. Spottiswoode and Amanda R. Ridley

Article citation details
Proc. R. Soc. B **287**: 20201140.
http://dx.doi.org/10.1098/rspb.2020.1140

Review timeline

Original submission:	13 January 2020
1st revised submission:	19 May 2020
2nd revised submission:	13 July 2020
Final acceptance:	13 July 2020

Note: Reports are unedited and appear as submitted by the referee. The review history appears in chronological order.

Review History

RSPB-2020-0072.R0 (Original submission)

Review form: Reviewer 1

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Excellent

Quality of the paper: Is the overall quality of the paper suitable?
Acceptable

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
Yes

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
Reviewer comments
This is an interesting study of a 15-year dataset on a cooperative-breeding bird living in an arid environment. The authors set out a clear hypothesis and use several approaches to test. Although they have a null result regarding their initial hypothesis, they still present interesting findings in relation to effects of heat stress and resources effects through rainfall on survival at different developmental stages. Overall, this work is well written and the introduction and discussion were clear and easy to follow. However, I found the different analyses done quite hard to follow and could at least by written, and perhaps done, in a more coherent way.

Major comments
A nice, clear set up and hypothesis and generally I found the introduction well written. But then I got lost in the different analyses you did – it looks like you follow different early stages of 'survival' but then start talking about survival generally and then mass gets thrown in there. I'm sure it fits together but it needs a clearer framework. This manuscript would benefit greatly from a clearer introduction to the different analyses and maybe a bit of re-structuring of the methods to make sure its clear what analysis you did and how they link together. E.g. l 142-144 you are talking about only survival and mass but then previously you talk about lots of stage of survival (l 71-72) and I found it hard to follow. I found the discussion well written and clearly set out, and it really brought your different results together an a coherent way. As mentioned below, I found it strange that you link this to negative effects of climate change, but don't test for trends in your survival estimates, only the rainfall and temperature covariates which do show interesting/worrying trends. It would be a stronger message if this was linked to declining fledgling survival for instance.

Methods
I have found the methods incredibly hard to follow. I don't really understand why you didn't you do one path analysis including survival of all development stages and the effect of nestling mass and environmental covariates on each stage? What is the added benefit of looking at overall survival when you have separated out each stage? And then have a separate path analysis looking at mass effects on overall survival. Unless I have misunderstood something, you could simply include all parameters in one path analysis and find the best fitting model of each survival state.
Also this makes the presentation of the analysis is seriously and at least more clarification on the model selection procedure is needed – just saying 'we checked whether terms crossed zero' is not enough and doesn't relate to the SI where there you mention AIC model selection of survival.

From what I can see you used an AIC approach for the Cox hazard model of 'overall survival' probabilities, but then just checked if they overlapped zero for the path analysis…? Why didn't you use AIC for the path analysis too- this is certainly possible. Renaming the sub-headings might help and you often refer to 'survival' but its not clear when you are talking about overall survival or survival of a particular stage.

Minor comments
Sometimes you say development stages and sometimes developmental stages- just make sure you're consistent
l. 44: ' Higher rainfall is often associated with improved reproductive success' I guess that only applies in arid regions, not clear in this sentence. And in fact throughout this paragraph, its written like it applies across ecosystems but only applies to arid zones.

l. 54 replace reproduction with reproductive?

l. 68 do you need '('pied babblers')' its already written once on l. 67

l. 74-77. so i guess the point is you need an interaction between temperature and group size effect- otherwise you cannot really claim that group size buffers negative temperature effects. You kind of state it but in a rather round-about way.

l. 87 would be helpful to mention here when the breeding season is and when each data for each parameter is collected. Otherwise its hard to follow the methods.

l.88 Is this necessary in the main text? Maybe put it in the acknowledgements/SI?

l. 98 you don't mention testing for possible trends in the temperature or rainfall covariates (and your demographic rates for the matter)? But then they come up in the results (l. 158). Since you have trends in covariates it may be worth detrending the data, to be sure they aren't just spurious correlations. Then it doesn't seem you test for trends in your demographic data – why not?
l. 118-119 'Model terms with HR confidence intervals not intersecting one were considered to explain significant patterns within our data.' But you have an AIC model selection section in the SI? As mentioned – your approach(es?) to model selection is confusing.

l. 126-128 just wondering why you didn't use a more common AIC-type model selection for the path analysis too? I don't get how you checked the significance of each term – by including all of them and then testing if they crossed zero? But this is totally dependent on the other terms included. Need some justification/more explanation here.

l. 130-131 (c) so it wasn't possible to include 'fledgling id' as a random effect here? I guess there are multiple fledglings per nest?

L 131-132 did you check for collinearity in covariates? I couldn't find anything about that.

L 137-138 its not an SEM– it's a confirmatory path analysis (as you say in the SI) when you have random effects included (e.g. Shipley, 2009, Ecology). Important to make the distinction since they are quite different procedures.

l. 197 'temperature' is quite vague- which of you T variables are you referring to ?

l. 283 did you consider looking at T variance rather than maximum? Just a thought, since you mention the importance of T extremes, perhaps group sizes is more important in buffering against variability.

l. 292 it does feel strange that you talk about future risks of climate change – but didn't check for any temporal trends in your different survival estimates? Is there any evidence of declines in the

demographic estimates and therefore an indication that we should be actually worried about climate change at this life history stage? Would be nice to see the annual estimates somewhere anyway.

Figure 1 – why haven't you included some confidence intervals on these predictions?

Figure 2 – same as figure 1, why have you only included CIs on the rainfall graph?

Figure S4 Maybe you don't need to show the non-significant coefficients? This would make the figure clearer (which is otherwise a very nice figure). Its almost a shame this is in an SI since I found it rather informative.

Review form: Reviewer 2

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Acceptable

General interest: Is the paper of sufficient general interest?
Acceptable

Quality of the paper: Is the overall quality of the paper suitable?
Acceptable

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
N/A

Is it clear?
N/A

Is it adequate?
N/A

Do you have any ethical concerns with this paper?
No

Comments to the Author
This is potentially an important paper and certainly makes use of an impressive data set. However, I found the paper to lack sufficient rationale, and to be based on a number of often disconnected and superficial analyses, which in my opinion need to be addressed.

First: For example, it is not clear why we should care whether studies split analyses down into developmental stages (Lines 39-40). Nor is it true to say that other studies have not considered buffering (see e.g. Covas et al. 2008; Rubenstein 2011). Finally, for the rationale, I don't understand how helpers might be expected to mitigate the effects of high temperature. Can helpers shade the brood, maybe I missed this, but if there is no rationale for helpers mitigating the effects of high temperatures, why include in the analyses interactions between helper number and temperature? The introduction needs more background information about the system and more rationale about what we expect before I can make sense of the results.

Second: There is a lot of attention made on Wiley & Ridley on the cut-off temperature effects and again, there is a whole section in the methods and results here on cut-offs, but there is no connection between the cut-off effects and helper effects. Surely, if you are going to show cut-offs, which is interesting, then we expect helper effects to mitigate the cut-offs, but these are not analysed, as far as I can tell. Finally, few climatic effects have been chosen for analyses, eg max temperature, but it is unclear why any one of the parameters was chosen, the time period over which they were chosen or validity. Perhaps the number of days over a given temperature would be more appropriate than max temperature, I don't know. But either way, I would like to see more evidence of a full effort being made to provide evidence for buffering, before it can be ruled out.

Figures: Given the rationale of this paper on buffering, I would have thought that all figures should include helper number interactions with weather parameters? Otherwise you are left with a series of non-novel figures.

Decision letter (RSPB-2020-0072.R0)

29-Feb-2020

Dear Ms Bourne:

I am writing to inform you that your manuscript RSPB-2020-0072 entitled "High temperatures drive offspring mortality in a cooperatively breeding bird" has, in its current form, been rejected for publication in Proceedings B.

This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

1) A 'response to referees' document including details of how you have responded to the comments, and the adjustments you have made.
2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
3) Line numbers in your main document.

To upload a resubmitted manuscript, log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Sincerely,
Dr Sasha Dall
mailto: proceedingsb@royalsociety.org

Associate Editor
Board Member: 1
Comments to Author:
The two reviewers and I all found this manuscript to be very interesting and timely, and the data set to be impressive. However, I concur with both reviewers that the statistical analysis is poorly explained and very hard to interpret. As currently presented, the statistical analyses are presented as several different methods, lacking a clear explanation about the overall goals of each analysis, how the analyses fit together, and why so many different methods were used. This makes it difficult to follow the rationale for each analysis and hard to understand how the results fit together. I agree with Reviewer 1's suggestion to clarify (and hopefully simplify?) the analyses. Second, please give general statistical information in the main text rather than the supplement (line 109); the format of Proc B allows such important information to be given along with the analyses rather than in a separate document. The results section would be easier to follow if the sub-sections were organized by finding (e.g. "Temperature influences survival") rather than by method (e.g. "Path analysis").

Given these concerns, I feel that the manuscript as currently presented is not publishable in Proceedings B, but a revised submission could be if the analyses were substantially overhauled and presented differently. You may also wish to consider moving some of the supplementary information (such as figure S4) to the main text since Reviewer 1 found it helpful.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
Reviewer comments
This is an interesting study of a 15-year dataset on a cooperative-breeding bird living in an arid environment. The authors set out a clear hypothesis and use several approaches to test. Although they have a null result regarding their initial hypothesis, they still present interesting findings in relation to effects of heat stress and resources effects through rainfall on survival at different developmental stages. Overall, this work is well written and the introduction and discussion were clear and easy to follow. However, I found the different analyses done quite hard to follow and could at least by written, and perhaps done, in a more coherent way.

Major comments
A nice, clear set up and hypothesis and generally I found the introduction well written. But then I got lost in the different analyses you did – it looks like you follow different early stages of

'survival' but then start talking about survival generally and then mass gets thrown in there. I'm sure it fits together but it needs a clearer framework. This manuscript would benefit greatly from a clearer introduction to the different analyses and maybe a bit of re-structuring of the methods to make sure its clear what analysis you did and how they link together. E.g. l 142-144 you are talking about only survival and mass but then previously you talk about lots of stage of survival (l 71-72) and I found it hard to follow. I found the discussion well written and clearly set out, and it really brought your different results together an a coherent way. As mentioned below, I found it strange that you link this to negative effects of climate change, but don't test for trends in your survival estimates, only the rainfall and temperature covariates which do show interesting/worrying trends. It would be a stronger message if this was linked to declining fledgling survival for instance.

Methods
I have found the methods incredibly hard to follow. I don't really understand why you didn't you do one path analysis including survival of all development stages and the effect of nestling mass and environmental covariates on each stage? What is the added benefit of looking at overall survival when you have separated out each stage? And then have a separate path analysis looking at mass effects on overall survival. Unless I have misunderstood something, you could simply include all parameters in one path analysis and find the best fitting model of each survival state.
Also this makes the presentation of the analysis is seriously and at least more clarification on the model selection procedure is needed – just saying 'we checked whether terms crossed zero' is not enough and doesn't relate to the SI where there you mention AIC model selection of survival. From what I can see you used an AIC approach for the Cox hazard model of 'overall survival' probabilities, but then just checked if they overlapped zero for the path analysis…? Why didn't you use AIC for the path analysis too- this is certainly possible. Renaming the sub-headings might help and you often refer to 'survival' but its not clear when you are talking about overall survival or survival of a particular stage.

Minor comments
Sometimes you say development stages and sometimes developmental stages- just make sure you're consistent

l. 44: ' Higher rainfall is often associated with improved reproductive success' I guess that only applies in arid regions, not clear in this sentence. And in fact throughout this paragraph, its written like it applies across ecosystems but only applies to arid zones.

l. 54 replace reproduction with reproductive?

l. 68 do you need '('pied babblers')' its already written once on l. 67

l. 74-77. so i guess the point is you need an interaction between temperature and group size effect- otherwise you cannot really claim that group size buffers negative temperature effects. You kind of state it but in a rather round-about way.

l. 87 would be helpful to mention here when the breeding season is and when each data for each parameter is collected. Otherwise its hard to follow the methods.

l.88 Is this necessary in the main text? Maybe put it in the acknowledgements/SI?

l. 98 you don't mention testing for possible trends in the temperature or rainfall covariates (and your demographic rates for the matter)? But then they come up in the results (l. 158). Since you have trends in covariates it may be worth detrending the data, to be sure they aren't just spurious correlations. Then it doesn't seem you test for trends in your demographic data – why not?

l. 118-119 'Model terms with HR confidence intervals not intersecting one were considered to explain significant patterns within our data.' But you have an AIC model selection section in the SI? As mentioned – your approach(es?) to model selection is confusing.

l. 126-128 just wondering why you didn't use a more common AIC-type model selection for the path analysis too? I don't get how you checked the significance of each term – by including all of them and then testing if they crossed zero? But this is totally dependent on the other terms included. Need some justification/more explanation here.

l. 130-131 (c) so it wasn't possible to include 'fledgling id' as a random effect here? I guess there are multiple fledglings per nest?

L 131-132 did you check for collinearity in covariates? I couldn't find anything about that.

L 137-138 its not an SEM– it's a confirmatory path analysis (as you say in the SI) when you have random effects included (e.g. Shipley, 2009, Ecology). Important to make the distinction since they are quite different procedures.

l. 197 'temperature' is quite vague- which of you T variables are you referring to ?

l. 283 did you consider looking at T variance rather than maximum? Just a thought, since you mention the importance of T extremes, perhaps group sizes is more important in buffering against variability.

l. 292 it does feel strange that you talk about future risks of climate change – but didn't check for any temporal trends in your different survival estimates? Is there any evidence of declines in the demographic estimates and therefore an indication that we should be actually worried about climate change at this life history stage? Would be nice to see the annual estimates somewhere anyway.

Figure 1 – why haven't you included some confidence intervals on these predictions?

Figure 2 – same as figure 1, why have you only included CIs on the rainfall graph?

Figure S4 Maybe you don't need to show the non-significant coefficients? This would make the figure clearer (which is otherwise a very nice figure). Its almost a shame this is in an SI since I found it rather informative.

Referee: 2

Comments to the Author(s)

This is potentially an important paper and certainly makes use of an impressive data set. However, I found the paper to lack sufficient rationale, and to be based on a number of often disconnected and superficial analyses, which in my opinion need to be addressed.

First: For example, it is not clear why we should care whether studies split analyses down into developmental stages (Lines 39-40). Nor is it true to say that other studies have not considered buffering (see e.g. Covas et al. 2008; Rubenstein 2011). Finally, for the rationale, I don't understand how helpers might be expected to mitigate the effects of high temperature. Can helpers shade the brood, maybe I missed this, but if there is no rationale for helpers mitigating the effects of high temperatures, why include in the analyses interactions between helper number and temperature? The introduction needs more background information about the system and more rationale about what we expect before I can make sense of the results.

Second: There is a lot of attention made on Wiley & Ridley on the cut-off temperature effects and again, there is a whole section in the methods and results here on cut-offs, but there is no connection between the cut-off effects and helper effects. Surely, if you are going to show cut-offs, which is interesting, then we expect helper effects to mitigate the cut-offs, but these are not analysed, as far as I can tell. Finally, few climatic effects have been chosen for analyses, eg max temperature, but it is unclear why any one of the parameters was chosen, the time period over which they were chosen or validity. Perhaps the number of days over a given temperature would be more appropriate than max temperature, I don't know. But either way, I would like to see more evidence of a full effort being made to provide evidence for buffering, before it can be ruled out.

Figures: Given the rationale of this paper on buffering, I would have thought that all figures should include helper number interactions with weather parameters? Otherwise you are left with a series of non-novel figures.

Author's Response to Decision Letter for (RSPB-2020-0072.R0)

See Appendix A.

RSPB-2020-1140.R0

Review form: Reviewer 1

Recommendation
Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
It was a pleasure to read through this manuscript again, the authors have thoroughly addressed the points raised by myself and reviewer 2. As a result, I found the manuscript and, specifically, the methods much easier to follow. I don't have any remaining major comments, but some minor ones are listed below.
The Authors now clearly set out the theory and their hypotheses surrounding the potentially mediating effects of group size, and even though they did not find such effects of group size, I think this is an interesting and relevant paper regarding the demographic impacts of climate change in arid environments. I do, however, get the feeling that - maybe as a consequence of your null result regarding the group size*env interaction - you play this point down throughout the manuscript. But as reviewer 2 mentioned before this is the novel point here - rather than the impacts of temperature extremes per se - and it's important to maintain focus on it and discuss possible explanations for why you did not find a significant effect (see my comment).
Also, it would help very much to clarify what you mean by reproduction success. Otherwise it's still unclear what exactly you're referring to, you talk about survival probabilities at each development stage and it's not explained whether that is the same as reproductive success. So please make it clear!
Specific comments
l.15-18. I still find these hypothesis confusing, without reading them a few times. What is the difference between what you refer to as survival if young (in i) vs. reproduction success in (ii)? Just make it as clear as possible how they fit together since its key to understanding what you did here.
l. 20. comma after 38degrees C
l. 54-57 strange use of semi-colons…
l. 65-76. I'm not very familiar with the literature on this but, belating to reviewer 2's points- in extreme dry/hot periods I wonder whether indeed this relationship you describe with group size still holds. I mean there will be very few resources available, and more (older) individuals to feed in large groups so perhaps older birds 'forgo' successful reproduction and rather survive (which I guess is the more important vital rate from a population-level perspective). At least might be worth putting in the discussion.
l. 85. Why is there a reference at the end of sentence setting out your hypotheses?
l. 83-88. Great, very clear now!
l.105. 'unique combination of metal and colour rings' – if it's similar as is commonly done, don't the metal rings have a unique id # and colours have a unique combination? Please make sure this statement is clear/correct.
l. 116. 2-3 not two to three. Or at least just be consistent, you use a mixture of numbers and letters elsewhere.
l. 123 I would write mean first and range after.
l. 151 write version not v.

l. 153-154. Just because two variables are < or > 2 doesn't mean they suddenly aren't or are correlated. Please make these sentences more nuanced.
l. 154. Remove: 'Sample sizes reflect data sets after removing records containing missing values.' Should be obvious I would say.

l. 179 MuMIn not MuMin.

l. 199. only

l. 224-229. Not mention of your key hypothesis here regarding the group size interaction?

l. 350-351. Could you at least add a sentence on the mechanism behind this group size nestling mass survival to fledging result.

l. 357. 'Lack of buffering effect of group size' might be more accurate.

Figure 1. Perhaps put the rainfall/temp time series (e) before figures a-d, I think it makes more sense.

Figure 3. Path analysis diagram looks great – could you make it a clearer somehow (maybe with circles rather than boxes or colours) which are the endogenous rather than exogenous variables in this model.

Review form: Reviewer 2

Recommendation
Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
N/A

Is it clear?
N/A

Is it adequate?
N/A

Do you have any ethical concerns with this paper?
No

Comments to the Author
I found the MS much improved. Personally, I am not convinced by the argument that helpers can buffer against high temperatures through load lightening. This is not a problem, as I might be wrong, but I at least want to highlight this because I do think some tempering of language is required in places. For example, I think you should be very careful not to overstretch your results. I believe you have found convincing evidence that helpers cannot buffer against high temperatures, but not that helpers cannot buffer period. You should therefore reduce generalisations made at the end of the Abstract and Discussion in particular. Further, I agree that high temperatures could cause a problem for desert species, but other options are available, like breed earlier in the year. I think you could present this as an option at the end of the Discussion. Otherwise, this is a nice study.

Decision letter (RSPB-2020-1140.R0)

29-Jun-2020

Dear Ms Bourne

I am pleased to inform you that your manuscript RSPB-2020-1140 entitled "High temperatures drive offspring mortality in a cooperatively breeding bird" has been accepted for publication in Proceedings B.

The referee(s) have recommended publication, but also suggest some minor revisions to your manuscript. Therefore, I invite you to respond to the referee(s)' comments and revise your manuscript. 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. If you do not think you will be able to meet this date please let us know.

To revise your manuscript, log into https://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote 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 referee(s) and upload a file "Response to Referees". You can use this to document any changes you make to the original manuscript. We require a copy of the manuscript with revisions made since the previous version marked as ‘tracked changes’ to be included in the ‘response to referees’ document.

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

1) A text file of the manuscript (doc, txt, rtf or tex), including the references, tables (including captions) and figure captions. Please remove any tracked changes from the text before submission. PDF files are not an accepted format for the "Main Document".

2) A separate electronic file of each figure (tiff, EPS or print-quality PDF preferred). The format should be produced directly from original creation package, or original software format. PowerPoint files are not accepted.

3) Electronic supplementary material: this should be contained in a separate file and where possible, all ESM should be combined into a single file. All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

4) A media summary: a short non-technical summary (up to 100 words) of the key findings/importance of your manuscript.

5) Data accessibility section and data citation
It is a condition of publication that data supporting your paper are made available either in the electronic supplementary material or through an appropriate repository.

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should be fully cited. To ensure archived data are available to readers, authors should include a 'data accessibility' section immediately after the acknowledgements section. This should list the database and accession number for all data from the article that has been made publicly available, for instance:
• DNA sequences: Genbank accessions F234391-F234402
• Phylogenetic data: TreeBASE accession number S9123
• Final DNA sequence assembly uploaded as online supplemental material
• Climate data and MaxEnt input files: Dryad doi:10.5521/dryad.12311
NB. From April 1 2013, peer reviewed articles based on research funded wholly or partly by RCUK must include, if applicable, a statement on how the underlying research materials – such as data, samples or models – can be accessed. This statement should be included in the data accessibility section.

If you wish to submit your data to Dryad (http://datadryad.org/) and have not already done so you can submit your data via this link
http://datadryad.org/submit?journalID=RSPB&manu=(Document not available) which will take you to your unique entry in the Dryad repository. If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link. Please see https://royalsociety.org/journals/ethics-policies/data-sharing-mining/ for more details.

6) For more information on our Licence to Publish, Open Access, Cover images and Media summaries, please visit https://royalsociety.org/journals/authors/author-guidelines/.

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

Sincerely,
Dr Sasha Dall
mailto: proceedingsb@royalsociety.org

Associate Editor
Board Member
Comments to Author:
Dear colleague
as you will see from the detailed comments provided, both reviewers are happy with your responses to their original queries and criticisms raised. Please make sure you carefully go through the remaining issues, in particular with respect to not over-generalising your results, e.g. that there is limited benefit beyond a certain temperature, and only in respect to the measured outcomes, rather than saying that there is no benefit at all.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).
It was a pleasure to read through this manuscript again, the authors have thoroughly addressed the points raised by myself and reviewer 2. As a result, I found the manuscript and, specifically, the methods much easier to follow. I don't have any remaining major comments, but some minor ones are listed below.
The Authors now clearly set out the theory and their hypotheses surrounding the potentially mediating effects of group size, and even though they did not find such effects of group size, I think this is an interesting and relevant paper regarding the demographic impacts of climate change in arid environments. I do, however, get the feeling that - maybe as a consequence of your null result regarding the group size*env interaction - you play this point down throughout the manuscript. But as reviewer 2 mentioned before this is the novel point here - rather than the impacts of temperature extremes per se - and it's important to maintain focus on it and discuss possible explanations for why you did not find a significant effect (see my comment).
Also, it would help very much to clarify what you mean by reproduction success. Otherwise it's still unclear what exactly you're referring to, you talk about survival probabilities at each development stage and it's not explained whether that is the same as reproductive success. So please make it clear!
Specific comments
l.15-18. I still find these hypothesis confusing, without reading them a few times. What is the difference between what you refer to as survival if young (in i) vs. reproduction success in (ii)? Just make it as clear as possible how they fit together since its key to understanding what you did here.
l. 20. comma after 38degrees C
l. 54-57 strange use of semi-colons…
l. 65-76. I'm not very familiar with the literature on this but, belating to reviewer 2's points- in extreme dry/hot periods I wonder whether indeed this relationship you describe with group size still holds. I mean there will be very few resources available, and more (older) individuals to feed in large groups so perhaps older birds 'forgo' successful reproduction and rather survive (which I guess is the more important vital rate from a population-level perspective). At least might be worth putting in the discussion.
l. 85. Why is there a reference at the end of sentence setting out your hypotheses?
l. 83-88. Great, very clear now!
l.105. 'unique combination of metal and colour rings' – if it's similar as is commonly done, don't the metal rings have a unique id # and colours have a unique combination? Please make sure this statement is clear/correct.

l. 116. 2-3 not two to three. Or at least just be consistent, you use a mixture of numbers and letters elsewhere.
l. 123 I would write mean first and range after.
l. 151 write version not v.
l. 153-154. Just because two variables are < or > 2 doesn't mean they suddenly aren't or are correlated. Please make these sentences more nuanced.
l. 154. Remove: 'Sample sizes reflect data sets after removing records containing missing values.' Should be obvious I would say.

l. 179 MuMIn not MuMin.

l. 199. only

l. 224-229. Not mention of your key hypothesis here regarding the group size interaction?

l. 350-351. Could you at least add a sentence on the mechanism behind this group size nestling mass survival to fledging result.

l. 357. 'Lack of buffering effect of group size' might be more accurate.

Figure 1. Perhaps put the rainfall/temp time series (e) before figures a-d, I think it makes more sense.

Figure 3. Path analysis diagram looks great – could you make it a clearer somehow (maybe with circles rather than boxes or colours) which are the endogenous rather than exogenous variables in this model.

Referee: 2

Comments to the Author(s).
I found the MS much improved. Personally, I am not convinced by the argument that helpers can buffer against high temperatures through load lightening. This is not a problem, as I might be wrong, but I at least want to highlight this because I do think some tempering of language is required in places. For example, I think you should be very careful not to overstretch your results. I believe you have found convincing evidence that helpers cannot buffer against high temperatures, but not that helpers cannot buffer period. You should therefore reduce generalisations made at the end of the Abstract and Discussion in particular. Further, I agree that high temperatures could cause a problem for desert species, but other options are available, like breed earlier in the year. I think you could present this as an option at the end of the Discussion. Otherwise, this is a nice study.

Author's Response to Decision Letter for (RSPB-2020-1140.R0)

See Appendix B.

Decision letter (RSPB-2020-1140.R1)

13-Jul-2020

Dear Ms Bourne

I am pleased to inform you that your manuscript entitled "High temperatures drive offspring mortality in a cooperatively breeding bird" has been accepted for publication in Proceedings B.

You can expect to receive a proof of your article from our Production office in due course, please check your spam filter if you do not receive it. PLEASE NOTE: you will be given the exact page length of your paper which may be different from the estimation from Editorial and you may be asked to reduce your paper if it goes over the 10 page limit.

If you are likely to be away from e-mail contact please let us know. Due to rapid publication and an extremely tight schedule, if comments are not received, we may publish the paper as it stands.

If you have any queries regarding the production of your final article or the publication date please contact procb_proofs@royalsociety.org

Your article has been estimated as being 10 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Open Access
You are invited to opt for Open Access, making your freely available to all as soon as it is ready for publication under a CCBY licence. Our article processing charge for Open Access is £1700. Corresponding authors from member institutions (http://royalsocietypublishing.org/site/librarians/allmembers.xhtml) receive a 25% discount to these charges. For more information please visit http://royalsocietypublishing.org/open-access.

Paper charges
An e-mail request for payment of any related charges will be sent out shortly. The preferred payment method is by credit card; however, other payment options are available.

Electronic supplementary material:
All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

You are allowed to post any version of your manuscript on a personal website, repository or preprint server. However, the work remains under media embargo and you should not discuss it with the press until the date of publication. Please visit https://royalsociety.org/journals/ethics-policies/media-embargo for more information.

Thank you for your fine contribution. On behalf of the Editors of the Proceedings B, we look forward to your continued contributions to the Journal.

Sincerely,
Editor, Proceedings B
mailto: proceedingsb@royalsociety.org

Appendix A

FitzPatrick Institute of **African Ornithology**
DST-NRF Centre of Excellence • **University of Cape Town**

Amanda Bourne
Room 2.18, John Day Zoology Building, UCT, 7701
Tel. (+27) (0)21 650-3306
Cell (+27) (0)83 440 2081
http://www.fitzpatrick.uct.ac.za/fitz/students/doctoral/bourne
email: abourne.uct@gmail.com

16 May 2020

Response to review

Dear Dr Dall and the Proceedings B review team

Thank you for the very comprehensive and thoughtful review of our manuscript, *High temperatures drive offspring mortality in a cooperatively breeding bird*, submitted on 13 January 2020. We appreciate the detailed feedback, fast response, and the opportunity to submit a revised manuscript. We have prepared a revision for resubmission, taking into account the comments we received from the associate editor and two reviewers on 29 February 2020. In this letter, we describe how we have addressed the comments in our revised text (our responses in *blue italics* below).

Please note that the line numbers refer to the 'clean' version of the revised manuscript, although we also include a copy of the original manuscript with all changes tracked.

Associate Editor

Comments to Author:
The two reviewers and I all found this manuscript to be very interesting and timely, and the data set to be impressive. However, I concur with both reviewers that the statistical analysis is poorly explained and very hard to interpret. As currently presented, the statistical analyses are presented as several different methods, lacking a clear explanation about the overall goals of each analysis, how the analyses fit together, and why so many different methods were used. This makes it difficult to follow the rationale for each analysis and hard to understand how the results fit together. I agree with Reviewer 1's suggestion to clarify (and hopefully simplify?) the analyses. Second, please give general statistical information in the main text rather than the supplement (line 109); the format of Proc B allows such important information to be given along with the analyses rather than in a separate document. The results section would be easier to follow if the sub-sections were organized by finding (e.g. "Temperature influences survival") rather than by method (e.g. "Path analysis").

Given these concerns, I feel that the manuscript as currently presented is not publishable in Proceedings B, but a revised submission could be if the analyses were substantially overhauled and presented differently. You may also wish to consider moving some of the supplementary information (such as figure S4) to the main text since Reviewer 1 found it helpful.

Thank you for the constructive feedback and the opportunity to revise our manuscript for resubmission. We are pleased that you found the manuscript interesting and timely. We have worked through the methods and results sections in some detail taking all of the feedback into account. Specifically, we have i) moved the overall survival analysis to the supplementary materials, as we agree this analysis was not adding much beyond a general context; ii) reanalysed the nest and fledgling survival data per development stage using only the separate GLMMs per development stage; iii) linked the segmented regression analyses used to identify temperature thresholds directly to the related GLMM data analyses for the same development stages; and iv) brought the path analysis for fledgling survival into the main text along with several sections of statistical methods that were previously in the supplementary material. We now present all the data and analyses for nests during incubation, nests during the chick-rearing period, and fledglings between fledge and independence together, rather than splitting them up by analytical approach as we had previously, and we analyse them all at the same scale. We hope that this helps to clarify and simplify the approach throughout. We have also reformatted the manuscript in the Proc B referencing style and included reference to several relevant new papers that have been published since we submitted our manuscript.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
Reviewer comments
This is an interesting study of a 15-year dataset on a cooperative-breeding bird living in an arid environment. The authors set out a clear hypothesis and use several approaches to test. Although they have a null result regarding their initial hypothesis, they still present interesting findings in relation to effects of heat stress and resources effects through rainfall on survival at different developmental stages. Overall, this work is well written and the introduction and discussion were clear and easy to follow. However, I found the different analyses done quite hard to follow and could at least by written, and perhaps done, in a more coherent way.

We are glad you enjoyed the manuscript and appreciate your feedback on the methods and results sections. Following your constructive suggestions, we have substantially revised the methods and results sections, and the analyses, in order to improve clarity and consistency; see reference to specific changes below:

Major comments
A nice, clear set up and hypothesis and generally I found the introduction well written. But then I got

lost in the different analyses you did – it looks like you follow different early stages of 'survival' but then start talking about survival generally and then mass gets thrown in there. I'm sure it fits together but it needs a clearer framework. This manuscript would benefit greatly from a clearer introduction to the different analyses and maybe a bit of re-structuring of the methods to make sure its clear what analysis you did and how they link together. E.g. l 142-144 you are talking about only survival and mass but then previously you talk about lots of stage of survival (l 71-72) and I found it hard to follow. I found the discussion well written and clearly set out, and it really brought your different results together a coherent way. As mentioned below, I found it strange that you link this to negative effects of climate change, but don't test for trends in your survival estimates, only the rainfall and temperature covariates which do show interesting/worrying trends. It would be a stronger message if this was linked to declining fledgling survival for instance.

Thanks for this great suggestion to include trends in our survival estimates! We have added trends for survival estimates into the analyses as suggested and show that there has been a trend of declining numbers of fledged nests and surviving fledglings over time in the study population. See Figure 1, and Lines 236-239: 'Both the number of nests fledged (a non-significant trend; $F_{1,12} = 3.747$, $p = 0.077$; Fig.1c) and the number of surviving young produced ($F_{1,12} = 5,285$, $p = 0.040$; Fig.1d) have declined at the study site since 2005, despite the number of groups monitored remaining relatively constant between years (coefficient of variation = 0.17).'

We have substantially restructured the methods and results section to explain more clearly why we used certain analyses and to simplify the methods used. We have moved the overall survival analysis from the main text to the supplementary materials (Fig S1) and cite Ridley (2016) to highlight that survival probabilities are not constant throughout early development in pied babblers. See lines 168-170: 'Pied babbler survival probabilities are not constant across time during early development [57], see Fig. S1, and covariates are unlikely to have the same relationship with survival during all three early development stages.'

The fact that survival probabilities are not constant throughout early development suggests there are different pressures during different early development stages (i.e. high temperatures during incubation and nestling stage, rainfall during the dependent fledgling period), which is why we wanted to analyse survival separately for each development stage. In this version of the manuscript, we move straight to the separate analyses as we agree with you that the overall survival analysis wasn't adding additional information and was confusing given the focus on different development stages in the introduction and discussion. GLMM analyses on survival per development stage are now presented alongside the segmented regressions on survival per development stage so that drivers of survival and the identification of temperature thresholds are dealt with simultaneously in the methods and results.

These analyses now focus consistently on survival at the level of the clutch or brood (i.e. at least one egg hatched, at least one chick fledged, at least one fledgling survived).

We have brought the path analysis back into the main text from the supporting information. We now explain the purpose of this analysis more clearly: Lines 207-213: "In addition to survival data at the scale of the breeding attempt, we have detailed individual-level survival data for 372 fledglings weighed and banded as 11-day-old nestlings. Larger nestling mass is commonly associated with higher survival probabilities in birds [16,77]. Prior research on pied babblers has shown that nestling mass is influenced by environmental factors such as temperature and rainfall [48]. We therefore used a confirmatory path analysis [78,79] to test for indirect effects of environmental and group size factors on survival to nutritional independence in known individual fledglings mediated via their mass as a nestling ($Mass_{11}$)."

We hope that this helps to clarify which analyses are important and how they fit together.

Methods

I have found the methods incredibly hard to follow. I don't really understand why you didn't you do one path analysis including survival of all development stages and the effect of nestling mass and environmental covariates on each stage? What is the added benefit of looking at overall survival when you have separated out each stage?

We agree that the overall survival analysis wasn't adding additional information and was confusing given the focus on different development stages in the introduction and discussion. Consequently, we have moved the overall survival analysis to the supplementary materials.

And then have a separate path analysis looking at mass effects on overall survival. Unless I have misunderstood something, you could simply include all parameters in one path analysis and find the best fitting model of each survival state.

We were unfortunately not able to run a single path analysis for all the parameters and development stages because we only measured nestling mass on day 11. To assess survival from one life stage to the next we use separate GLMMs at the brood level and then constructed one path analysis at the level of the individual fledgling for the subset of the data where we had additional information on moderating variables (mass). However, your observation did lead us to an inconsistency in our analytical approach, which we have now corrected. Previously we analysed the progression from egg to nestling and from nestling to fledgling at the scale of the clutch or brood (because we did not have specific information about each individual egg or nestling) and the progression from fledgling to nutritionally independent juvenile at the scale of the individual fledgling (allowing the inclusion of nestling mass in the GLMM analyses). We now conduct all the GLMM analyses at the scale of the clutch or brood, and use the path analysis to test for mediated effects of environmental and social factors on survival to nutritional independence via effects on nestling mass at the individual level. While it is possible to rank different path analysis models by AIC our goal was not to choose between

competing hypotheses, and but rather to construct a single model testing the relative importance of direct effects of environmental and group size factors vs. effects mediated via nestling mass. We hope that our approach is much clearer now, in both the methods and results sections. See Lines 207-229 and Lines 279-300.

Also this makes the presentation of the analysis is seriously and at least more clarification on the model selection procedure is needed – just saying 'we checked whether terms crossed zero' is not enough and doesn't relate to the SI where there you mention AIC model selection of survival. From what I can see you used an AIC approach for the Cox hazard model of 'overall survival' probabilities, but then just checked if they overlapped zero for the path analysis…? Why didn't you use AIC for the path analysis too- this is certainly possible. Renaming the sub-headings might help and you often refer to 'survival' but its not clear when you are talking about overall survival or survival of a particular stage.

We have substantially reworked both the methods and results sections in line with your feedback, and reanalysed all of the data. We now clarify which model selection and significance assessment processes apply to which analyses in the main text, moving these out of the supplementary material for greater clarity. While it is possible to rank different candidate path analysis models by AIC, our goal was not to choose between competing hypotheses, or to compare different path models, but rather to construct a single model testing the extent to which environmental and social factors influence survival to nutritional independence directly, or indirectly via their effects on nestling mass. We have renamed all the sub-headings in the results section to more clearly speak to the theme of each section, highlighting the findings rather than analytical approach. All analyses presented in the main text now refer to survival within a particular development stage, consistent with the hypotheses described in the introduction.

Minor comments

Sometimes you say development stages and sometimes developmental stages- just make sure you're consistent

Thanks for picking this up – corrected throughout. We went with development rather than developmental.

l. 44: ' Higher rainfall is often associated with improved reproductive success' I guess that only applies in arid regions, not clear in this sentence. And in fact throughout this paragraph, its written like it applies across ecosystems but only applies to arid zones.

Adjusted as follows: Line 47-48: 'For birds in arid environments, higher rainfall is often associated with improved reproductive success [21,22], …'

l. 54 replace reproduction with reproductive?

The structure of the paragraph has changed to the extent that this edit no longer applies.

l. 68 do you need '('pied babblers')' its already written once on l. 67

Adjusted as follows to clarify that this sets the tone for how the species will be referred to throughout the rest of the manuscript: Line 77-78: 'We use a comprehensive 15-year dataset on southern pied babblers Turdoides bicolor (hereafter 'pied babblers'),'

l. 74-77. so i guess the point is you need an interaction between temperature and group size effect- otherwise you cannot really claim that group size buffers negative temperature effects. You kind of state it but in a rather round-about way.

Adjusted as follows: Line 85-88: 'If the presence of helpers buffers the effect of environmental variation on reproduction, as proposed by the temporal variability hypothesis [9], then we would expect an interaction between environmental factors and group size, such that weaker impacts of adverse climatic conditions on reproduction are observed in larger groups.'

l. 87 would be helpful to mention here when the breeding season is and when each data for each parameter is collected. Otherwise its hard to follow the methods.

Adjusted as follows: Line 98-99: 'They breed during the austral summer, from September to March [57].'; Line 110-111: 'Data were collected for each austral summer breeding season from September 2005–February 2019 (14 breeding seasons in total).

We also now include a lot more detail on when and how each parameter was collected (see methods Line 112-149). These explanations were previously located in the supporting information.

l.88 Is this necessary in the main text? Maybe put it in the acknowledgements/SI?

We have moved the animal ethics statement to the acknowledgements as suggested

l. 98 you don't mention testing for possible trends in the temperature or rainfall covariates (and your demographic rates for the matter)? But then they come up in the results (l. 158). Since you have trends in covariates it may be worth detrending the data, to be sure they aren't just spurious correlations. Then it doesn't seem you test for trends in your demographic data – why not?

Thanks for this. We have now added that we test for trends in the temp and rainfall covariates and added tests for reproduction trends into the methods:

Line 162-165: 'We tested for temporal trends in environmental (temperature and rainfall) and reproductive (nest success, fledgling survival) parameters using univariate linear models with breeding season as the only predictor.'

And we present the results of these analyses in our new Fig. 1 and Lines 232-242, as mentioned above.

Based on your suggestion, we have detrended the data, using the 'detrend' function in the R package 'pracma', and explain this in the methods section:

Line 165-166: Covariates exhibiting temporal trends were detrended using the detrend *function in the package* pracma *[71].*

l. 118-119 'Model terms with HR confidence intervals not intersecting one were considered to explain significant patterns within our data.' But you have an AIC model selection section in the SI? As mentioned – your approach(es?) to model selection is confusing.

The reference to model terms not intersecting one only refers to the Cox Proportional Hazards model for overall survival between initiation of incubation and nutritional independence. We have removed this analysis from main manuscript text and refer to it only for context in the supporting information, where the relevant methods are explained separately to avoid confusion with the other analyses. AIC model selection processes were used for the GLMM analyses for each development stage, which are our main focus. We have brought the model selection methods description into the main text now (Lines 165-188) so that it is clear that this approach refers to the GLMM analyses for each development stage.

l. 126-128 just wondering why you didn't use a more common AIC-type model selection for the path analysis too? I don't get how you checked the significance of each term – by including all of them and then testing if they crossed zero? But this is totally dependent on the other terms included. Need some justification/more explanation here.

We have used path analysis to test the support for a single specific hypothesis that we had based on knowledge of the study species: as we were not comparing different path models, an AICc approach (such that we used for the GLMMs where we were comparing different candidate models) is not valid here. We know from previous work on pied babblers that the same factors that influenced survival probabilities in our GLMM analyses also influence nestling mass. Therefore, we used the single path analysis model specifically to explore the relative importance of different pathways via which temperature, rainfall and group size could affect survival to independence: i.e. what proportion of that impact was direct, vs indirect, i.e. mediated by the effects of these variables on nestling mass and the subsequent influence of nestling mass on survival. We were able to do this for the fledge-to-independence development stage because we collect nestling mass measurements 11 days after hatching and so have data at the individual level for mass and survival during this development stage. We have clarified the purpose of this analysis, and the approach we have taken, on Lines 216-223:

'While model selection processes can be applied to multiple path analyses [81], our goal was not to choose between competing hypotheses, and but rather to construct a single model testing the relative importance of direct effects of environmental and group size factors vs. effects mediated via nestling

mass. Path analysis allowed us to specify and simultaneously quantify all hypothesised relationships of interest, including the indirect effects of weather and group size on survival via nestling mass. Path coefficients are partial regression coefficients and can be interpreted similarly to simple and multiple regression outputs. Statistical significance was taken as $p < 0.05$.'

l. 130-131 (c) so it wasn't possible to include 'fledgling id' as a random effect here? I guess there are multiple fledglings per nest?

Analyses are at the scale of the breeding attempt, so brood identity is the unit of analysis. We did not include information about individual fledglings in these analyses as they were only uniquely identifiable after ringing, 11 days after hatching. Clarified Lines 172-174: 'These analyses were undertaken at the level of the breeding attempt (i.e. clutch or brood) because individual offspring were only ringed for individual identification from the 11th day after hatching, by which time ~60% of monitored breeding attempts had failed.' and Lines 183-186: 'We considered the influence of the following parameters on (a) the probability of at least one egg per clutch surviving to hatch, (b) the probability of at least one nestling per brood surviving to fledge, and (c) the probability of at least one fledgling per brood surviving to nutritional independence'. Our random term was group identity.

L 131-132 did you check for collinearity in covariates? I couldn't find anything about that.

Yes, Lines 152-155: 'All explanatory variables were tested for correlation with one another [67]. Mean T_{max90} and mean T_{min90} were correlated (VIF = 2.8) and all other explanatory variables were not correlated with each other (all VIF < 2). Correlated variables were not included in same additive models.'

L 137-138 its not an SEM– it's a confirmatory path analysis (as you say in the SI) when you have random effects included (e.g. Shipley, 2009, Ecology). Important to make the distinction since they are quite different procedures.

Corrected as suggested on Lines 211-213: 'We therefore used a confirmatory path analysis [78,79] to test for indirect effects of environmental and group size factors on survival to nutritional independence in known individual fledglings mediated via their mass as a nestling ($Mass_{11}$)'

We now use the term path analysis rather than SEM throughout.

l. 197 'temperature' is quite vague- which of you T variables are you referring to ?

Clarified throughout.

l. 283 did you consider looking at T variance rather than maximum? Just a thought, since you mention the importance of T extremes, perhaps group sizes is more important in buffering against variability.

Thank you, yes, we considered other temperature variables as well, specifically average daily minimum temperatures and average daily temperature variation (T_{max}-T_{min}). Daily minimum temperatures and daily variation in temperature were not significant predictors of survival in any of our exploratory analyses but we now include them for transparency in the full glm models tested as shown in Tables S2-S4 in the Supplementary Materials. Also see main text methods:

Line 145-149: 'Daily minimum (T_{min}) and maxium (T_{max}) temperatures, daily temperature variation (T_{max} - T_{min}), were averaged for each development stage: incubation (mean T_{minInc}, mean T_{maxInc}, mean T_{varInc}), nestling (mean $T_{minBrood}$, mean $T_{maxBrood}$, mean $T_{varBrood}$), and fledgling (mean $T_{min90,}$ mean $T_{max90,}$ mean T_{var90}).'

Line 183-188: 'We considered the influence of the following parameters on (a) the probability of at least one egg per clutch surviving to hatch, (b) the probability of at least one nestling per brood surviving to fledge, and (c) the probability of at least one fledgling per brood surviving to nutritional independence: for (a) group size, $Rain_{60}$, mean T_{minInc}, mean T_{maxInc}, and mean T_{varInc}, for (b) group size, $Rain_{60}$, mean $T_{minBrood}$, mean $T_{varBrood}$, mean $T_{maxBrood}$, and mean $T_{maxBrood}^2$, and for (c) group size, $Rain_{90}$, mean $T_{min90,}$ mean $T_{var90,}$ mean $T_{max90,}$ and mean T_{max90}^2.'

l. 292 it does feel strange that you talk about future risks of climate change – but didn't check for any temporal trends in your different survival estimates? Is there any evidence of declines in the demographic estimates and therefore an indication that we should be actually worried about climate change at this life history stage? Would be nice to see the annual estimates somewhere anyway.
There certainly are! Thanks for this interesting suggestion; we have included evidence of declining numbers of fledglings and independent young as discussed above.

Figure 1 – why haven't you included some confidence intervals on these predictions?
Please note that figure 1 has changed during the revision and is no longer the same figure to which this comment refers. With the changes made to the manuscript and the other figures we are now presenting in the main text, we now present the estimates from the glmm models, with se and confidence interval, exclusively in Table 1.

Figure 2 – same as figure 1, why have you only included CIs on the rainfall graph?
Please note that figure 2 has changed during the revision and is no longer the same figure to which this comment refers. We now present the estimates from the glmm models, with se and confidence interval, in Table 1, see above comment.

Figure S4 Maybe you don't need to show the non-significant coefficients? This would make the figure clearer (which is otherwise a very nice figure). Its almost a shame this is in an SI since I found it rather informative.
We have brought the path analysis model figure back into the main text. All specified pathways are shown on the figure as is conventional for visualisation of path analyses.

Referee: 2

Comments to the Author(s)

This is potentially an important paper and certainly makes use of an impressive data set. However, I found the paper to lack sufficient rationale, and to be based on a number of often disconnected and superficial analyses, which in my opinion need to be addressed.

Thank you for your feedback; we appreciate your thorough engagement with the material and your comments have helped us to clarify the rationale throughout (see below) and to simplify and clarify our methods and results.

First: For example, it is not clear why we should care whether studies split analyses down into developmental stages (Lines 39-40). Nor is it true to say that other studies have not considered buffering (see e.g. Covas et al. 2008; Rubenstein 2011). Finally, for the rationale, I don't understand how helpers might be expected to mitigate the effects of high temperature. Can helpers shade the brood, maybe I missed this, but if there is no rationale for helpers mitigating the effects of high temperatures, why include in the analyses interactions between helper number and temperature? The introduction needs more background information about the system and more rationale about what we expect before I can make sense of the results.

We have now added our rationale for why we think it is worth considering temperature and also considering development stages separately:

Lines 37-43: 'To date, however, there are few empirical studies that explicitly test the extent to which group-living mitigates the effects of climate variability on reproduction [10–13]. Of these studies, only Langmore et al. (2016) and van de Ven et al. (2019) explore impacts of temperature alongside variation in group size, despite evidence of thermoregulatory benefits of group living [14], and only Covas et al. [13] consider offspring survival across more than one development stage. This latter point is important, because specific drivers of survival can differ substantially between development stages [15–17].

Thank you for suggesting the Covas et al.(2008) paper – we have found it very useful and have incorporated it. Rubenstein's recent paper (Guindre-Parker &Rubenstein 2020) is an example of an explicit test of buffering, and we now include reference to it. This latter paper had not yet been published when we first submitted this manuscript.

We have added more detail to the rationale for why we might expect helpers to mitigate the effects of high temperatures into the introduction, see above and also:

Lines 65-76: 'A likely mechanism underlying such benefits of cooperation is load-lightening [42], which refers to individual reductions in workload in response to the presence of additional group members. Load-lightening has been observed in a number of cooperatively-breeding species [11,16,43], and may operate via task-partitioning [37,44] or by improved access to resources [45,46]. In larger groups, there are more individuals available to assist with breeding attempts, which can either lead to load-lightening amongst individual group members [11,37], or to cumulatively greater investment in

young [36,51]; both are potential benefits of group living that may be particularly advantageous during unfavourable rainfall or temperature conditions. Specifically, these effects could mean that larger groups are better able to maintain adequate levels of parental care to eggs, nestlings, and/or fledglings at high temperatures or during periods of low rainfall, despite individual declines in investment in parental care behaviours.'
We have also added more background information about the study system into the methods section, see Lines 91-108.

Second: There is a lot of attention made on Wiley & Ridley on the cut-off temperature effects and again, there is a whole section in the methods and results here on cut-offs, but there is no connection between the cut-off effects and helper effects. Surely, if you are going to show cut-offs, which is interesting, then we expect helper effects to mitigate the cut-offs, but these are not analysed, as far as I can tell.
We agree that we have not directly connected helper effects with the precise temperature cutoff identified in this or previous work on southern pied babblers. The threshold identified by Wiley & Ridley (2016) and du Plessis et al (2012) has been mentioned in a minor, contextual section of the manuscript to illustrate a temporal trend of increasing temperature at the study site, see Results section, Lines 232-234: The total number of days (Oct–Mar) exceeding 35.5°C, identified as a critical temperature threshold in pied babblers [48,59], has increased significantly at the study site since 2005 ($F_{1,12}$ = 7.448, p = 0.018; Fig.1a)"
The range of important temperature thresholds that we have identified here are similar to the thresholds identified by du Plessis et al (2012) and Wiley & Ridley (2016) but vary somewhat between development stages, ranging from 35.4°C to 37.3°C.
For all of the cases where we show a significant negative effect of temperature on survival probabilities, we have demonstrated in the glmm analyses that there is no moderating effect of group size. We therefore do not expect that group size will influence the identified temperature thresholds.

We hope that the significant restructure of the methods and results sections will help to clarify these aspects.

Finally, few climatic effects have been chosen for analyses, eg max temperature, but it is unclear why any one of the parameters was chosen, the time period over which they were chosen or validity. Perhaps the number of days over a given temperature would be more appropriate than max temperature, I don't know. But either way, I would like to see more evidence of a full effort being made to provide evidence for buffering, before it can be ruled out.

Thank you, yes, we considered other temperature variables as well, specifically average daily minimum temperatures, average daily temperature variation (T_{max}-T_{min}), and the number and proportion of hot days for each time period. Please see our detailed response to Reviewer 1 on the inclusion of temperature variation and minimum temperature. Regards whether number of hot days would have been a better predictor, we tested proportion of hot days previously and found effects in the same direction but that mean daily maximum was consistently the better predictor by AIC:

Model	*AIC*	
	Mean Tmax	*Proportion Hot Days*
Survival to hatching	*602.4*	*625.3*
Survival to fledging	*459.6*	*462.1*
Survival to nutritional independence	*208.5*	*221.1*

Mean T_{max} over the breeding attempt as a predictor is also consistent with other recent studies considering effects of environmental factors on fitness in cooperative breeders, e.g. van de Ven et al (2019). While it would certainly be very interesting to further consider impacts of heatwaves, heatwave return rates, numbers of days above a threshold temperature or even number of hours above a threshold temperature (Cunningham et al. 2013a; Sharpe et al. 2019), we found Mean T_{max} informative and a powerful predictor across all development stages in our analyses, promoting internal consistency. We hope and intend to explore the other variables in future work.
Please see our discussion about the time periods of relevance for each analysis as these related to the temperature and rainfall variables used:
Line 145-149: "Daily minimum (T_{min}) and maxium (T_{max}) temperatures, daily temperature variation (T_{max} - T_{min}), were averaged for each development stage: incubation (mean T_{minInc}, mean T_{maxInc}, mean T_{varInc}), nestling (mean $T_{minBrood}$, mean $T_{maxBrood}$, mean $T_{varBrood}$), and fledgling (mean $T_{min90,}$ mean $T_{max90,}$ mean T_{var90}). Rainfall was summed for the 60 days prior to initiation of incubation ($Rain_{60}$), and for the period between fledging and independence ($Rain_{90}$)."

Figures: Given the rationale of this paper on buffering, I would have thought that all figures should include helper number interactions with weather parameters? Otherwise you are left with a series of non-novel figures.

In the substantial revision of the manuscript we have chosen to present a different set of figures – the climate and demographic trends, temperature thresholds of concern, and the path analysis testing the extent to which effects of temperature, rainfall and group size influence survival via their effects of nestling mass. Results of our models testing the interactions between helper numbers and weather parameters are now exclusively shown in the tables included in both the main text and the supplementary materials (Table 1, Tables S2-S4).

Appendix B

Running head: Temperature drives offspring mortality

Title: High temperatures drive offspring mortality in a cooperatively breeding bird

Response to Referees: July 2020

Dear Dr Dall

Thank you for facilitating the review of our manuscript 'High temperatures drive offspring mortality in a cooperatively breeding bird'. We are thrilled to hear that the referees found the manuscript improved and have accepted it for publication in Proceedings B. Please see below (in blue) our responses outlining how we have addressed the remaining comments. We have also attached a copy of the manuscript with track changes showing what was changed.

Warm regards

Amanda Bourne

Fitzpatrick Institute of African Ornithology

Abourne.uct@gmail.com

Associate Editor
Board Member
Comments to Author:

Dear colleague
as you will see from the detailed comments provided, both reviewers are happy with your responses to their original queries and criticisms raised. Please make sure you carefully go through the remaining issues, in particular with respect to not over-generalising your results, e.g. that there is limited benefit beyond a certain temperature, and only in respect to the measured outcomes, rather than saying that there is no benefit at all.

Thank you very much for your role in facilitating this review and for your helpful feedback. We have made some adjustments in the abstract and discussion to avoid over-generalising our results – please see our responses to specific responses in this regard below, along with our responses to the other points raised by the two reviewers.

We have also made minor changes to the text, figures, and supplementary materials in order to conform to the page limit requirements at Proceedings B.

Please let me know if there is anything further I can assist with as we prepare the manuscript for publication at Proceedings B.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).

It was a pleasure to read through this manuscript again, the authors have thoroughly addressed the points raised by myself and reviewer 2. As a result, I found the manuscript and, specifically, the methods much easier to follow. I don't have any remaining major comments, but some minor ones are listed below.
The Authors now clearly set out the theory and their hypotheses surrounding the potentially mediating effects of group size, and even though they did not find such effects of group size, I think this is an interesting and relevant paper regarding the demographic impacts of climate change in arid environments. I do, however, get the feeling that - maybe as a consequence of your null result regarding the group size*env interaction - you play this point down throughout the manuscript. But as reviewer 2 mentioned before this is the novel point here - rather than the impacts of temperature extremes per se - and it's important to maintain focus on it and discuss possible explanations for why you did not find a significant effect (see my comment).
Also, it would help very much to clarify what you mean by reproduction success. Otherwise it's still unclear what exactly you're referring to, you talk about survival probabilities at each development stage and it's not explained whether that is the same as reproductive success. So please make it clear!

Thank you for reviewing our manuscript for a second time and providing your feedback. We value your input, and are pleased that you have enjoyed the manuscript, finding our research question interesting. We now clarify throughout the manuscript that we have studied survival probabilities of young and, rather than trying to differentiate offspring survival from reproductive success, we now only refer to what we measured specifically – being the survival of young as eggs, chicks, and fledglings.

Thanks also for your comment about maintaining the focus on our research question regarding the group size * envir interaction. We agree that this is the most novel aspect of our study and have attempted to give this aspect plenty of coverage in the abstract, introduction, and conclusion. Given the null result (which we also think is extremely interesting) and strict page limits for the manuscript, along with the requests from reviewer 2 and the editor to avoid over-generalising from this null result, we do give slightly greater prominence to the significant relationships in the results and discussion, which include enviro and social factors, but not the interaction between them.

Specific comments
l.15-18. I still find these hypothesis confusing, without reading them a few times. What is the difference between what you refer to as survival if young (in i) vs. reproduction success in (ii)? Just make it as clear as possible how they fit together since its key to understanding what you did here.

We have rewritten the relevant sections of the abstract to clarify and report only what we actually measured (offspring survival rather than reproductive success)

Line 14-22: "We use a 15-year dataset on a cooperative-breeding arid-zone bird, the southern pied babbler *Turdoides bicolor*, to test i) whether environmental conditions and group size correlate with survival of young during three development stages (egg, nestling, fledgling), and ii) whether group size mitigates the impacts of adverse environmental conditions on survival of young. Exposure to high mean daily maximum temperatures (mean T_{max}) during early development was associated with reduced survival probabilities of young in all three

development stages. No young survived when mean T_{max} > 38°C, across all group sizes. Low survival of young at high temperatures has broad implications for recruitment and population persistence in avian communities given the rapid pace of advancing climate change"

l. 20. comma after 38degrees C

Added the comma as suggested

l. 54-57 strange use of semi-colons…

Changed the semi-colons to commas

l. 65-76. I'm not very familiar with the literature on this but, belating to reviewer 2's points- in extreme dry/hot periods I wonder whether indeed this relationship you describe with group size still holds. I mean there will be very few resources available, and more (older) individuals to feed in large groups so perhaps older birds 'forgo' successful reproduction and rather survive (which I guess is the more important vital rate from a population-level perspective). At least might be worth putting in the discussion.

We address this possibility in the discussion:

Line 347-349: "This suggests that physiological tolerance limits [97] and resource constraints [98] at high temperatures may exceed any potential buffering effect of group size on offspring survival in cooperative breeders in arid and semi-arid environments [10]."

Line 351-354: "In this study, negative effects of adverse climate conditions on breeding success in a cooperative breeder were not moderated by group size, suggesting that reproduction in pied babblers is constrained by available resources and physiology at high temperatures and low rainfall, regardless of group size."

l. 85. Why is there a reference at the end of sentence setting out your hypotheses?

Removed the reference. We describe the use of 90 days as a biologically relevant cut-off age, and source the relevant study, on lines 120-121.

l. 83-88. Great, very clear now!

Thank you

l.105. 'unique combination of metal and colour rings' – if it's similar as is commonly done, don't the metal rings have a unique id # and colours have a unique combination? Please make sure this statement is clear/correct.

We have clarified the difference between the metal and colour rings and that we use one metal ring and up to three colour rings:

Line 97-99: "Birds in the study population are marked as nestlings with a metal band (engraved with a unique number) and a unique combination of up to three colour rings for individual identification"

l. 116. 2-3 not two to three. Or at least just be consistent, you use a mixture of numbers and letters elsewhere.

Changed to 2-3 as suggested, for consistency with the previous paragraph

l. 123 I would write mean first and range after.

Moved mean first and range afterwards for both instances

l. 151 write version not v.

We have written out version in full as suggested

l. 153-154. Just because two variables are < or > 2 doesn't mean they suddenly aren't or are correlated. Please make these sentences more nuanced.

We have now added the correlation coefficients:

Line 142-144: "Mean T_{max90} and mean T_{min90} were correlated (VIF = 2.8, correlation coefficient = 0.77) and all other explanatory variables were not correlated with each other (all VIF < 1.3, correlation coefficients < 0.30)."

l. 154. Remove: 'Sample sizes reflect data sets after removing records containing missing values.' Should be obvious I would say.

Removed the sentence as suggested

l. 179 MuMIn not MuMin.

Changed to MuMIn

l. 199. Only

Corrected – thanks for picking up this typo

l. 224-229. Not mention of your key hypothesis here regarding the group size interaction?

We had already tested the enviro*group size interaction in the GLMMs and found that the interaction term was not a significant predictor of fledgling survival. Including interaction terms in the path analysis only showed no significant effect of the interaction ($p > 0.351$), complicating the path analysis without adding new information. The new information that the path analysis adds is a test of whether the observed effects of environment and group size on survival of at least one fledgling per brood was mediated via their mass as an 11-day-old-nestling, using a dataset we had for individual young weighed at 11 days old.

l. 350-351. Could you at least add a sentence on the mechanism behind this group size ◊ nestling mass ◊ survival to fledging result.

We have added a sentence on the possible mechanism:

Line 336-337: "Larger nestlings are more mobile and better developed at fledging, enabling them to forage more effectively, avoid predators, and survive longer [94]"

l. 357. 'Lack of buffering effect of group size' might be more accurate.

Line 342-343: "We found a lack of a buffering effect of group size on the effects of high mean T_{max} on offspring survival."

Figure 1. Perhaps put the rainfall/temp time series (e) before figures a-d, I think it makes more sense.

Due to space constraints we now only show the time series panel in the main text of the manuscript and have moved the temp, rainfall, and breeding trend figures to the supplementary materials. We have restructured the results paragraph accordingly.

Line 2i8-228: "Most rain falls between Dec and Feb (72%), when temperatures are high (Fig.1). Most pied babbler breeding activity occurs between Oct and Dec (68%), when conditions are generally drier and cooler than later in the season (Fig.1). The total number of days (Oct–Mar) exceeding 35.5°C, identified as a critical temperature threshold in pied babblers [53,54], has increased significantly at the study site since 2005 ($F_{1,12}$ = 7.448, p = 0.018; Fig.S3a). Total summer rainfall (Oct–Mar) over the same time period was highly variable but showed a declining, non-statistically significant trend ($F_{1,12}$ = 1.616, p = 0.228; Fig.S3b). Both the number of nests fledged (a non-significant trend; $F_{1,12}$ = 3.747, p = 0.077; Fig.S3c) and the number of surviving young produced ($F_{1,12}$ = 5,285, p = 0.040; Fig.S3d) have declined at the study site since 2005, despite the number of groups monitored remaining relatively constant between years (coefficient of variation = 0.17)."

Figure 3. Path analysis diagram looks great – could you make it a clearer somehow (maybe with circles rather than boxes or colours) which are the endogenous rather than exogenous variables in this model.

Thank you. We have differentiated the response variables within the model using a light orange shading and updated the figure legend accordingly.

Line 645-464: "R2 for component models are given (grey shaded boxes) above response variables (orange shaded boxes)."

Referee: 2

Comments to the Author(s).
I found the MS much improved. Personally, I am not convinced by the argument that helpers can buffer against high temperatures through load lightening. This is not a problem, as I might be wrong, but I at least want to highlight this because I do think some tempering of language is required in places. For example, I think you should be very careful not to overstretch your results. I believe you have found convincing evidence that helpers cannot buffer against high temperatures, but not that helpers cannot buffer period. You should therefore reduce generalisations made at the end of the Abstract and Discussion in particular. Further, I agree that high temperatures could cause a problem for desert species, but other options are available, like breed earlier in the year. I think you could present this as an option at the end of the Discussion. Otherwise, this is a nice study.

Thank you for your feedback, and for taking the time to review our revised manuscript. Following your advice, we have tempered some of the generalisations made in the Abstract and Discussion. We agree with you that our results do not suggest that cooperative breeding is never advantageous, simply that the availability of more helpers does not appear to buffer against the effects of high temperatures on reproductive outcomes. In this study, we show that "larger group sizes were indirectly associated with increased survival via the positive effect

of larger group size on nestling $Mass_{11}$." (Lines 264-265) and "group size influenced individual fledgling survival probabilities indirectly, via a positive effect on nestling mass." (Lines 334-335). We also acknowledge the work of previous studies on the benefits of cooperation, including in pied babblers: "Benefits of cooperation include earlier fledging age and more broods raised per season [31], reduced costs of breeding for females [11,32], enhanced egg investment [33], increased fledgling recruitment [17,34], and the ability to raise overlapping broods [12,35]" (line 50-52).

We do not conclude that there is no effect of helpers per se, but we are saying specifically in the context of buffering environmental effects, there is no evidence for a group size*enviro interaction.

With regards tempering the language around the lack of buffering effect, please see the adjusted text below:

From the end of the Abstract (line 22-25):

"Impacts of high temperatures on survival of young were not moderated by group size, suggesting that the availability of more helpers in a group is unlikely to buffer against compromised offspring survival as average and maximum temperatures increase with rapid anthropogenic climate change."

And in the Discussion

Line 347-349: The lack of an observed buffering effect "physiological tolerance limits [97] and resource constraints [98] at high temperatures may exceed any potential buffering effect of group size on offspring survival in cooperative breeders in arid and semi-arid environments [10]."

Line 363-367: "Our findings suggest that the presence of more helpers in a group is unlikely to provide a buffer against reproductive failure as average and maximum temperatures increase with advancing anthropogenic climate change and raise concerns for the long-term persistence of arid zone species in the face of rapidly changing environmental conditions."

We also now present breeding earlier in the season as a potential way that pied babblers might mitigate the effects of high temperatures on breeding attempts:

Line 313-317: "This could undermine population growth and ultimately lead to local extinctions for this species, although the effect may be mitigated by behavioural adjustments such as breeding earlier in the season or engaging in compensatory breeding during good years [59]."