Why men invest in non-biological offspring: paternal care and paternity confidence among Himba pastoralists

Sean P. Prall and Brooke A. Scelza

Review timeline

Original submission:	17 September 2019	Note: Reports are unedited and appear as
1st revised submission:	11 December 2019	submitted by the referee. The review history
2nd revised submission:	20 January 2020	appears in chronological order.
Final acceptance:	18 February 2020	

Review History

RSPB-2019-2165.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?
Acceptable

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

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?
Yes

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
The reference claming that the population studied displays high extra paternity is not published. Thus the claim is difficult to evaluate. In addition, the proportion of extra-paternity is not reported here…
How many Omoka and non-omoka individuals were analyzed ? A descriptive table of the sample is missing (e.g. the number of observation could only be found in tables from the supplementary materials). Is the proportion of Omoka and non-omoka individuals analyzed here different from the rate of non-paternity reported (in the unpublished paper) ?

« All coefficient results and posterior predictions show 89% prediction intervals, to avoid confusion with significance tests. » This is odd. The justification to show 89% prediction intervals is non-scientific. Why not show 95 % prediction intervals, and add in the figure legend something like : « bars show 95 % prediction intervals » ? (people do known the difference between a confidence and a prediction interval).

Lines 289. WAIC values are indicated as "very similar". How much is similar, and when do you consider that it is not similar ? Generally, no quantitative description of the results is made. For example "Curves that fall further from zero demonstrate a larger effect of omoka status for that sex.". What is a "larger" effect ? How further is considered acceptable to decide that it is worth reporting as an effect ?

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?
Excellent

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
The present manuscript provides data on Himba pastoralist men's views and patterns of paternal care contingent upon a child's status as the product of an affair (omoka). The study involves vignettes among a convenience sample of 16 Himba, and data on four outcomes—children's anthropometrics as measures of growth; children's fostering status; fathers' livestock gifts; and fathers' brideprice—to test whether fathers' investments are contingent upon the paternity status of children. The study builds on a rich body of quantitative and culturally contextualized sexuality and parenting research among the Himba by the investigators. The manuscript is well-written and polished, making it easy to follow. The Figures are also excellent visual representations of key findings. There are few such rich multi-methods human case studies on topics like this suited for the journal, which is part of what makes this a strong contribution. That said, a number of concerns can also be raised, as I do below.
A key methods and empirical concern is the lack of clear reporting of sample size and sample demographics. How many fathers' data are represented for key analyses? What is the age range among those fathers, and how might they vary across other key variables such as economic background, living arrangements, fertility, etc.? Information about sample size is essential and needs to be clear, as this also speaks to concerns such as whether adequate statistical power exists to test all predictions. Moreover, one might wonder about cohort changes in expectations of and practices by fathers, with such concerns amplified the greater the age range of fathers. If some men have more children, all else being equal, they may have less to devote to specific children. How do the Himba assess omoka status? How well does Himba-assessed omoka status map on to genetic paternity? Please include these key aspects of paternity. There is also some conceptual clarity needed about the paper's aim: is it getting at "paternity confidence" (title) or paternity (which seems to be the claim with analyses contingent upon omoka status)? Extending these thoughts, it would be helpful to hear how Himba men themselves understand biological paternity, since it's implicit in the manuscript that they do this accurately. Perhaps underscore somewhere, too, that despite the availability of genetic paternity data here, all analyses are

predicated on distinctions of Himba self-reported omoka status rather than genetically-based paternity.

In the Introduction when discussing the theoretical significance of paternity, it might make good foundational theoretical sense to cite Hamilton (1964) on kin selection. Around references 36/37, it might also make foundational sense to explicitly note and reference the concept that male parenting effort may be mating effort, which is related to the points made by the authors about signaling, reputation, etc.

Please provide some summary rationale for the cohesion of the four outcomes (e.g., children's anthropometrics, brideprice) evaluated here, and note why other potential outcomes from theory and the literature (e.g., children's survival, socioemotional capacities, female age of menarche, children's age of reproductive landmarks such as first reproduction) were not evaluated. I don't take issue with these four measures employed having ethnographic salience, but that I think there could be more clear rationale for why these and not others.

The literature discusses a variety of factors that can shape paternal investment. One is sex ratio, including Schact's work in Guyana. What is the sex ratio in this Himba community and more broadly (i.e., Himba region), and could that be related to the prevalence of practices (infidelity, nonpaternity) that are remarkably high compared to other human populations? How often are Himba men away from the household (non-residence) or community, due to work or other obligations, and could this be a relevant factor for inferring non-paternity (child omoka status) plus the normative allowance for infidelity? Scelza's work on Himba infidelity has noted that infidelity is higher among arranged marriages compared to love matches. Does this marital relationship variable (arranged vs. love) influence the findings reported in the manuscript, or for that matter any other key measure of marital dynamics (like monogamous vs. polygynous marriage)?

For the livestock gifts measure, over what time frame was this measured? One-time gift? Lifetime? Last year?

I think brideprice was measured as an amount given for a son's first wife; is that correct? What fraction (100%?) is the son's brideprice is paid by the focal father? I am wondering about potential confounders here. Is there any evidence that omoka men might be married to "lower cost" brides, like Borgerhoff Mulder's findings that wealthier men could pay more for younger wives with higher reproductive value? Are any other family members kicking in amounts toward brideprice to offset or in response to the amount offered by the focal father?

I liked the vignettes as a method to evaluate paternal expectations depending upon child's omoka status. Given that the findings were from a convenience sample of 16 men, I did wonder if a comparable sample of women would give comparable responses or not. That would likely be a different manuscript, though from a sexual conflict model perspective, that could be an interesting topic to investigate.

As another speculative inquiry that could be ignored if relevant data are not available, what do Himba fathers say when a child dies? This manuscript is devoid of emotional dynamics (which is fine given its aims), though a complementary approach would explore fathers' and children's sentiments regarding these father-child relationships (e.g., varied outcomes assessed in Gray and Brown 2015 St. Kitts fatherhood paper in "Fathering"). A genetic or social father's child's death might serve as an emotional barometer of the meaning of the child to the father. If methodologically tractable, it would be interesting to try quantifying whether genetic and non-genetic Himba fathers report comparable or different emotional attachments to these children, respectively. Regardless, in the scope of evolutionary approaches to human behavior, this Himba case study material on nonpaternity, including not just focusing on homicide (like in classic work by Daly and Wilson) as the key dependent variable, is a welcome contribution to the literature.

One final speculative topic, with perhaps some quick note of relevance in future directions, could be any changing patterns in fertility and children's work/education among the Himba. What is average fertility in the sample of fathers reported on here, and is that changing (e.g., lowering)? The ethnographic observations about girls and boys economic contributions in the Discussion were interesting, but one could imagine the costs/benefits to children to change (like Bock's work in Botswana) as child's labor and formal education also change, leaving one to wonder about potential for predicted Himba change too.

Line 349: the first letter of brideprice is the number 8 in my manuscript version.

Decision letter (RSPB-2019-2165.R0)

15-Oct-2019

Dear Dr Prall:

I am writing to inform you that your manuscript RSPB-2019-2165 entitled "Why men invest in non-biological offspring:
Paternal care and paternity confidence among Himba pastoralists" 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.

In your revision process, please take a second look at how open your science is; our policy is that all data involved with the study should be made openly accessible-- see:
https://royalsociety.org/journals/ethics-policies/data-sharing-mining/
Insufficient sharing of data can delay or even cause rejection of a paper.

Sincerely,
Professor John Hutchinson, Editor
mailto: proceedingsb@royalsociety.org

Associate Editor
Comments to Author:
Your paper presents a mixed methods approach to understand Himba pastoralists patterns of parental care, which is novel and interesting. However, I think the main findings (and the novelty

of these) could be made more explicit in order to engage a broad audience. Reviewer 2 is enthusiastic about the paper and offers many detailed comments to improve the manuscript, but also agrees with the main concerns of Reviewer 1 concerning the data treatment. The manuscript needs more information about your sample and how you have treated this statistically, and needs explicit justification for the approach you have taken.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
The reference claming that the population studied displays high extra paternity is not published. Thus the claim is difficult to evaluate. In addition, the proportion of extra-paternity is not reported here…
How many Omoka and non-omoka individuals were analyzed ? A descriptive table of the sample is missing (e.g. the number of observation could only be found in tables from the supplementary materials). Is the proportion of Omoka and non-omoka individuals analyzed here different from the rate of non-paternity reported (in the unpublished paper) ?

« All coefficient results and posterior predictions show 89% prediction intervals, to avoid confusion with significance tests. » This is odd. The justification to show 89% prediction intervals is non-scientific. Why not show 95 % prediction intervals, and add in the figure legend something like : « bars show 95 % prediction intervals » ? (people do known the difference between a confidence and a prediction interval).

Lines 289. WAIC values are indicated as "very similar". How much is similar, and when do you consider that it is not similar ? Generally, no quantitative description of the results is made. For example "Curves that fall further from zero demonstrate a larger effect of omoka status for that sex.". What is a "larger" effect ? How further is considered acceptable to decide that it is worth reporting as an effect ?

Referee: 2

Comments to the Author(s)
The present manuscript provides data on Himba pastoralist men's views and patterns of paternal care contingent upon a child's status as the product of an affair (omoka). The study involves vignettes among a convenience sample of 16 Himba, and data on four outcomes—children's anthropometrics as measures of growth; children's fostering status; fathers' livestock gifts; and fathers' brideprice—to test whether fathers' investments are contingent upon the paternity status of children. The study builds on a rich body of quantitative and culturally contextualized sexuality and parenting research among the Himba by the investigators. The manuscript is well-written and polished, making it easy to follow. The Figures are also excellent visual representations of key findings. There are few such rich multi-methods human case studies on topics like this suited for the journal, which is part of what makes this a strong contribution. That said, a number of concerns can also be raised, as I do below.
A key methods and empirical concern is the lack of clear reporting of sample size and sample demographics. How many fathers' data are represented for key analyses? What is the age range among those fathers, and how might they vary across other key variables such as economic background, living arrangements, fertility, etc.? Information about sample size is essential and needs to be clear, as this also speaks to concerns such as whether adequate statistical power exists to test all predictions. Moreover, one might wonder about cohort changes in expectations of and practices by fathers, with such concerns amplified the greater the age range of fathers. If some men have more children, all else being equal, they may have less to devote to specific children.

How do the Himba assess omoka status? How well does Himba-assessed omoka status map on to genetic paternity? Please include these key aspects of paternity. There is also some conceptual clarity needed about the paper's aim: is it getting at "paternity confidence" (title) or paternity (which seems to be the claim with analyses contingent upon omoka status)? Extending these thoughts, it would be helpful to hear how Himba men themselves understand biological paternity, since it's implicit in the manuscript that they do this accurately. Perhaps underscore somewhere, too, that despite the availability of genetic paternity data here, all analyses are predicated on distinctions of Himba self-reported omoka status rather than genetically-based paternity.

In the Introduction when discussing the theoretical significance of paternity, it might make good foundational theoretical sense to cite Hamilton (1964) on kin selection. Around references 36/37, it might also make foundational sense to explicitly note and reference the concept that male parenting effort may be mating effort, which is related to the points made by the authors about signaling, reputation, etc.

Please provide some summary rationale for the cohesion of the four outcomes (e.g., children's anthropometrics, brideprice) evaluated here, and note why other potential outcomes from theory and the literature (e.g., children's survival, socioemotional capacities, female age of menarche, children's age of reproductive landmarks such as first reproduction) were not evaluated. I don't take issue with these four measures employed having ethnographic salience, but that I think there could be more clear rationale for why these and not others.

The literature discusses a variety of factors that can shape paternal investment. One is sex ratio, including Schact's work in Guyana. What is the sex ratio in this Himba community and more broadly (i.e., Himba region), and could that be related to the prevalence of practices (infidelity, nonpaternity) that are remarkably high compared to other human populations? How often are Himba men away from the household (non-residence) or community, due to work or other obligations, and could this be a relevant factor for inferring non-paternity (child omoka status) plus the normative allowance for infidelity? Scelza's work on Himba infidelity has noted that infidelity is higher among arranged marriages compared to love matches. Does this marital relationship variable (arranged vs. love) influence the findings reported in the manuscript, or for that matter any other key measure of marital dynamics (like monogamous vs. polygynous marriage)?

For the livestock gifts measure, over what time frame was this measured? One-time gift? Lifetime? Last year?

I think brideprice was measured as an amount given for a son's first wife; is that correct? What fraction (100%?) is the son's brideprice is paid by the focal father? I am wondering about potential confounders here. Is there any evidence that omoka men might be married to "lower cost" brides, like Borgerhoff Mulder's findings that wealthier men could pay more for younger wives with higher reproductive value? Are any other family members kicking in amounts toward brideprice to offset or in response to the amount offered by the focal father?

I liked the vignettes as a method to evaluate paternal expectations depending upon child's omoka status. Given that the findings were from a convenience sample of 16 men, I did wonder if a comparable sample of women would give comparable responses or not. That would likely be a different manuscript, though from a sexual conflict model perspective, that could be an interesting topic to investigate.

As another speculative inquiry that could be ignored if relevant data are not available, what do Himba fathers say when a child dies? This manuscript is devoid of emotional dynamics (which is fine given its aims), though a complementary approach would explore fathers' and children's sentiments regarding these father-child relationships (e.g., varied outcomes assessed in Gray and Brown 2015 St. Kitts fatherhood paper in "Fathering"). A genetic or social father's child's death might serve as an emotional barometer of the meaning of the child to the father. If methodologically tractable, it would be interesting to try quantifying whether genetic and non-genetic Himba fathers report comparable or different emotional attachments to these children, respectively. Regardless, in the scope of evolutionary approaches to human behavior, this Himba case study material on nonpaternity, including not just focusing on homicide (like in classic work by Daly and Wilson) as the key dependent variable, is a welcome contribution to the literature.

One final speculative topic, with perhaps some quick note of relevance in future directions, could be any changing patterns in fertility and children's work/education among the Himba. What is average fertility in the sample of fathers reported on here, and is that changing (e.g., lowering)? The ethnographic observations about girls and boys economic contributions in the Discussion were interesting, but one could imagine the costs/benefits to children to change (like Bock's work in Botswana) as child's labor and formal education also change, leaving one to wonder about potential for predicted Himba change too.

Line 349: the first letter of brideprice is the number 8 in my manuscript version.

Author's Response to Decision Letter for (RSPB-2019-2165.R0)

See Appendix A.

RSPB-2019-2890.R0

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?
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?
Yes

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?
   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
As the authors decided to answer the comments from the reviewers in a separate file (which is unusual), it is not readily possible to evaluate their answers in direct relation with the points raised. After some copy/paste work, it appears that most points were considered, but some questions were not addressed. For example, the question "Is the proportion of Omoka and non-omoka individuals analyzed here different from the rate of non-paternity reported (in the unpublished paper) ?" is left unanswered. Some points require further thoughts.

A. The claim that "Fathers have been shown to be very accurate in their paternity assertions..." is now quantitatively substantiated: "..correctly assessing paternity 73% of the time" (line 139). Thus 27% of men are not correctly assessing their paternity status. I would not claim that this is "very accurate".

B. Lines 207-212: "Here we rely on paternity assertions rather than genetic paternity, as it is the beliefs of men about paternity that are most relevant to investment. However, as these data are from the same set of households used in the genetic paternity study, which showed high accuracy in detecting nonpaternity events, we are confident that there is a strong link between paternity assertions and genetic paternity."
That belief (paternity assertions) is stronger than genes (genetic paternity) for paternal investment is perhaps true. That does not mean that genes has no effect. Why not test both, as the data is already there ? Instead of being "confident that there is a strong link between paternity assertions and genetic paternity", please show this link (is this the 73% of correct assessment ??), as the data is already there (or cite a reference). Anyway, the fact that 27% of men have a wrong belief about their paternity is not a strong argument to justify to rely only on paternity assertions...

C. Again with the prediction interval. I maintain that the sentence (still in the revised version) stating that "All coefficient results and posterior predictions show 89% prediction intervals, to avoid confusion with significance tests." is not correct. I agree that any value (89%, 95%) is arbitrary, but the *justification* expressed here to use 89% is not acceptable to me. This is because, just to avoid a confusion, a different value is chosen, with consequences on results, graphical appearance, etc. A confusion could be avoided by another mean than choosing a different value (e.g. more explanations in the text). A different justification should be provided to stick to this value of 89%. Incidentally, a search on Google Scholar with "89% prediction interval" gives 3 results (yes, only three; to check, do not forget the quotes on the request). A similar search with 90% give 2590 results, and a search with "95% prediction interval" provides around 12000 results. Thus the claim (in the answers to reviewers) that the 89% is "common in modern Bayesian statistics" is not very strong. Please justify properly this choice of 89% which seems, indeed, very unusual in Bayesian statistics.

D. Where are the figure legends ?

Review form: Reviewer 3

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
Statistical Review
As asked, I will focus mostly on the statistical aspects of the paper.
I think I can clear up the question why 89% was chosen as a prediction interval: this number appears repeatedly in McElreath's 'Rethinking Statistics' textbook on Bayesian statistics, which is widely used in the social sciences. The author there specifically states that this number was chosen because it is a prime number and easily remembered, and also points out that it is useful to avoid 95% because readers otherwise unconsciously interpret results as though they were hypothesis tests. One thing the same author points out is that in the end, any value chosen here is up to the scientist, and that for normally distributed posteriors (as is the case here), using 80% or 89% or 90% or 95% will give nearly identical results. It would be good if the manuscript's authors could provide this reference in the text. Also, while they do not have to justify the number itself, it would be good to justify why they chose a relatively more relaxed prediction interval than is customary.
With regards to the results, I do not think that the choice of the prediction interval will influence them dramatically, given that the main result is anyways a null result (no difference in treatment). From the plots and estimates, it appears that there is solid evidence for the interpretation given, especially considering that anthropological data are fairly messy and the data sets fairly small and are thereby usually accompanied by a relatively large error.
It would help the reader if the graphs would include a representation of the actual data. It seems that the omoka throughout show much greater variation in estimates; it would be interesting to

see whether this is the result of some fathers treating their omokas really bad and others really good (bimodal distribution), or whether there is a continuum in how they are treated.

Decision letter (RSPB-2019-2890.R0)

14-Jan-2020

Dear Dr Prall:

Your manuscript has now been peer reviewed and the reviews have been assessed by an Associate Editor. The reviewers' comments (not including confidential comments to the Editor) and the comments from the Associate Editor are included at the end of this email for your reference. As you will see, the reviewers and the Editors have raised some concerns with your manuscript and we would like to invite you to revise your manuscript to address them. The general judgement is that the revisions are achievable and it is likely the main conclusions are robust enough.

We do not allow multiple rounds of revision so we urge you to make every effort to fully address all of the comments at this stage. If deemed necessary by the Associate Editor, your manuscript will be sent back to one or more of the original reviewers for assessment. If the original reviewers are not available we may invite new reviewers. Please note that we cannot guarantee eventual acceptance of your manuscript at this stage.

To submit your revision please 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 Revision". Your manuscript number has been appended to denote a revision.

When submitting your revision please upload a file under "Response to Referees" in the "File Upload" section. This should document, point by point, how you have responded to the reviewers' and Editors' comments, and the adjustments you have made to the 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.

Your main manuscript should be submitted as a text file (doc, txt, rtf or tex), not a PDF. Your figures should be submitted as separate files and not included within the main manuscript file.

When revising your manuscript you should also ensure that it adheres to our editorial policies (https://royalsociety.org/journals/ethics-policies/). You should pay particular attention to the following:

Research ethics:
If your study contains research on humans please ensure that you detail in the methods section whether you obtained ethical approval from your local research ethics committee and gained informed consent to participate from each of the participants.

Use of animals and field studies:
If your study uses animals please include details in the methods section of any approval and licences given to carry out the study and include full details of how animal welfare standards were ensured. Field studies should be conducted in accordance with local legislation; please include details of the appropriate permission and licences that you obtained to carry out the field work.

Data accessibility and data citation:
It is a condition of publication that you make available the data and research materials supporting the results in the article. Datasets should be deposited in an appropriate publicly available repository and details of the associated accession number, link or DOI to the datasets must be included in the Data Accessibility section of the article (https://royalsociety.org/journals/ethics-policies/data-sharing-mining/). Reference(s) to datasets should also be included in the reference list of the article with DOIs (where available).

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should also be fully cited and listed in the references.

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.

For more information please see our open data policy http://royalsocietypublishing.org/data-sharing.

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. Please try to submit all supplementary material as a single file.

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].

Please submit a copy of your revised paper within three weeks. If we do not hear from you within this time your manuscript will be rejected. If you are unable to meet this deadline please let us know as soon as possible, as we may be able to grant a short extension.

Thank you for submitting your manuscript to Proceedings B; we look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Best wishes,
Professor John Hutchinson, Editor
mailto: proceedingsb@royalsociety.org

Associate Editor Board Member
Comments to Author:
Reviewer 1 has some continuing concerns about the statistical analysis. The statistical reviewer has considered these problems and thinks some additional information needs to be added to make the choice of analysis clearer.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).

As the authors decided to answer the comments from the reviewers in a separate file (which is unusual), it is not readily possible to evaluate their answers in direct relation with the points raised. After some copy/paste work, it appears that most points were considered, but some questions were not addressed. For example, the question "Is the proportion of Omoka and non-omoka individuals analyzed here different from the rate of non-paternity reported (in the unpublished paper) ?" is left unanswered. Some points require further thoughts.

A. The claim that "Fathers have been shown to be very accurate in their paternity assertions..." is now quantitatively substantiated: "..correctly assessing paternity 73% of the time" (line 139). Thus 27% of men are not correctly assessing their paternity status. I would not claim that this is "very accurate".

B. Lines 207-212: "Here we rely on paternity assertions rather than genetic paternity, as it is the beliefs of men about paternity that are most relevant to investment. However, as these data are from the same set of households used in the genetic paternity study, which showed high accuracy in detecting nonpaternity events, we are confident that there is a strong link between paternity assertions and genetic paternity."
That belief (paternity assertions) is stronger than genes (genetic paternity) for paternal investment is perhaps true. That does not mean that genes has no effect. Why not test both, as the data is already there ? Instead of being "confident that there is a strong link between paternity assertions and genetic paternity", please show this link (is this the 73% of correct assessment ??), as the data is already there (or cite a reference). Anyway, the fact that 27% of men have a wrong belief about their paternity is not a strong argument to justify to rely only on paternity assertions...

C. Again with the prediction interval. I maintain that the sentence (still in the revised version) stating that "All coefficient results and posterior predictions show 89% prediction intervals, to avoid confusion with significance tests." is not correct. I agree that any value (89%, 95%) is arbitrary, but the *justification* expressed here to use 89% is not acceptable to me. This is because, just to avoid a confusion, a different value is chosen, with consequences on results, graphical appearance, etc. A confusion could be avoided by another mean than choosing a different value (e.g. more explanations in the text). A different justification should be provided to stick to this value of 89%. Incidentally, a search on Google Scholar with "89% prediction interval" gives 3 results (yes, only three; to check, do not forget the quotes on the request). A similar search with 90% give 2590 results, and a search with "95% prediction interval" provides around 12000 results. Thus the claim (in the answers to reviewers) that the 89% is "common in modern Bayesian statistics" is not very strong. Please justify properly this choice of 89% which seems, indeed, very unusual in Bayesian statistics.

D. Where are the figure legends ?

Referee: 3

Comments to the Author(s).
Statistical Review
As asked, I will focus mostly on the statistical aspects of the paper.
I think I can clear up the question why 89% was chosen as a prediction interval: this number appears repeatedly in McElreath's 'Rethinking Statistics' textbook on Bayesian statistics, which is widely used in the social sciences. The author there specifically states that this number was

chosen because it is a prime number and easily remembered, and also points out that it is useful to avoid 95% because readers otherwise unconsciously interpret results as though they were hypothesis tests. One thing the same author points out is that in the end, any value chosen here is up to the scientist, and that for normally distributed posteriors (as is the case here), using 80% or 89% or 90% or 95% will give nearly identical results. It would be good if the manuscript's authors could provide this reference in the text. Also, while they do not have to justify the number itself, it would be good to justify why they chose a relatively more relaxed prediction interval than is customary.

With regards to the results, I do not think that the choice of the prediction interval will influence them dramatically, given that the main result is anyways a null result (no difference in treatment). From the plots and estimates, it appears that there is solid evidence for the interpretation given, especially considering that anthropological data are fairly messy and the data sets fairly small and are thereby usually accompanied by a relatively large error.

It would help the reader if the graphs would include a representation of the actual data. It seems that the omoka throughout show much greater variation in estimates; it would be interesting to see whether this is the result of some fathers treating their omokas really bad and others really good (bimodal distribution), or whether there is a continuum in how they are treated.

Author's Response to Decision Letter for (RSPB-2019-2890.R0)

See Appendix B.

RSPB-2019-2890.R1 (Revision)

Review form: Reviewer 1

Recommendation
Accept as is

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

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
Ok, this is fine...

Decision letter (RSPB-2019-2890.R1)

18-Feb-2020

Dear Dr Prall

I am pleased to inform you that your manuscript entitled "Why men invest in non-biological offspring:
Paternal care and paternity confidence among Himba pastoralists" has been accepted for publication in Proceedings B. Congratulations!!

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

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.

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

Paper charges
An e-mail request for payment of any related charges will be sent out after proof stage (within approximately 2-6 weeks). 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.

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,
Professor John Hutchinson
Editor, Proceedings B
mailto: proceedingsb@royalsociety.org

Associate Editor:
Board Member: 1
Comments to Author:
The authors have done a good job in addressing all the final points and I think the paper will make a good contribution to the literature.

Appendix A

Response to Reviewers

Dear Dr. Hutchinson,

We submit for your consideration a resubmission of RSPB-2019-2165 entitled, "Paternal care and paternity confidence among Himba pastoralists." We thank the reviewers for their insightful comments, which we carefully considered as we revised our manuscript. We have made several significant changes to the manuscript to address your comments, and the comments of the reviewers. These include: (1) a more explicit treatment of the main results, highlighting their novelty and significance, (2) more detail on our choice of statistical tests and on the analyses themselves, (3) additional details on sample descriptives for the four datasets. Detailed responses to reviewer comments are below, in addition to track changes in the manuscript.

Reviewer 1
(1) *Publication of EPP Rate*. Our publication showing the rate of EPP in this population has now been published and we have updated the reference. It was under review at the time of our initial submission. We have also now included some more detail from that paper on the rate of EPP and the rate of paternity confidence. We believe this strengthens our case in this paper because we are now able to definitively show both that the risk of EPP is high in this population and that men are aware and quite accurate in knowing which children are theirs. Please also see our response to R2 (point 3) below.
(2) *Description of the Sample*: We have added a table (Table S1) of descriptive statistics about the sample, including the numbers of purported omoka and non-omoka children, as well as other relevant details. We had to divide this into four parts, as the samples used for each of the outcome variables were different (e.g. child anthropometrics vs data on brideprice which came from married adult sons). We have also included in the text (197-200, 207-212) an explanation of how the *omoka* assertions included here compare with the EPP and paternity confidence rates from the genetic data. We would also note that these rates should vary based on sample type, and exclusion criteria as described in the main text. For example, when looking at the effect of paternity confidence on anthropometrics, we exclude children who are known to be fostered out, since fosterage has negative and sex specific effects on child health and growth in this population (see Prall & Scelza 2017, Am J Hum Biol 29:6; Scelza & Silk 2014, Hum Nat 25:4).
(3) *Predicted Probabilities vs Confidence Intervals*: While the use of 95% confidence intervals is traditional, our choice of 89% prediction intervals is common when using Bayesian statistics. Confidence interval size (as well as many aspects of statistics, such as P values, etc) is arbitrary, and not based on any real theory. We use 89% intervals, to distinguish between the predictive power of the model and null-hypothesis significance testing. The intervals used here (89%), as well as others (90%, 96%, etc) are common in modern

Bayesian statistics, and indeed are standard with some software (see McElreath's 2015 *Statistical Rethinking*).

(4) *Interpreting WAIC Scores*: This is a valid point, as WAIC scores can be confusing to interpret. We have changed "very similar' to "nearly identical," to highlight that we are not making arbitrary decisions about size. This is a common mistake to many who use WAIC or other information criterion, particularly given that the information criterion themselves are estimates with an error about the mean. This uncertainty in WAIC is often unreported, and the slightest differences in WAIC used to support hypotheses. These nuances are not appropriate for the main body of our paper, given that we aren't drawing strong conclusions based on model comparison. However, we would refer the reviewer to the supplementary materials, which contain detailed reporting for all models, including WAIC comparison and model weight. We have been intentionally brief in reporting these results, since there are too many models to consider (three for each anthropometric outcome, for example), and instead prefer to report these largely null results outside of the main manuscript. The reviewer also raises the issue of effect size as shown in the difference in posteriors plots on several of the figures. These plots show the probability distribution of differences between o*moka* and non-*omoka* individuals on the standardized residual scale. So size of difference here is a meaningful unit of value, and the probability distribution illustrates the confidence in these estimates. To aid in interpretation, we report probabilities of the difference being non-zero whenever results are reported. This type of reporting is increasingly common in Bayesian analysis.

Reviewer 2

(1) *Description of the Sample*: See response to R1 and the addition of Table S1.

(2) *Cohort Effects*: Generally, we agree with R1 that the quantity-quality trade-off is likely at play (i.e. fathers with more children will have less time to devote to each one) and this is actually critical to our point, in that we want to see who men favor when resources are limited. The reviewer also raises important issues about sample size.. As requested we have included more details about sample size (see Figure S1, and have also included all data and code with this submission). With respect to issues of power, such concerns are always at play regarding work in small-scale populations, and particularly true when considering the nature of the questions addressed in this manuscript. We strive to be as transparent as possible regarding the potential limitations of this study, but the results presented here are largely unambiguous (both in effect of predictors, difference in posteriors, and model comparison metrics). The Bayesian nature of our analysis means we are also very conservative in our assumption of any effect of our predictors (using highly regularized priors, indicating no prior assumption of any effect, and so not biasing analysis to any assumed outcome). If indeed much larger sample sizes (such in the brideprice analysis) may ultimately yield meaningful differences, the results presented here suggest that, at best, effect size is very small. To the reviewer's point about whether there are

generational differences in the way that men thinking about paternity and parenting, we would point to our norms study, which shows near unanimity in responses to the questions about whether bias against *omoka* children is ok, despite a very large age range in this sample(range of 19 to 79, mean=42.7, SD= 18.8). However, we acknowledge this as a limitation and have added a paragraph to this effect to the Discussion (lines 499-506).

(3) *Paternity Assessments*: We have included additional ethnographic data on the ways in which men assess paternity, and also refer to the now published genetic data we have from this population on EPP and paternity confidence, which we believe strengthens the manuscript, but which was under review at the time of our initial submission. We have also included some more explicit statements about our use of paternity assertions and linked those to the data we have on confidence to clarify this link. We also explain that we used paternity assertions (confidence) rather than genetic paternity, because we believe it is what men think about paternity that should most affect their investment decisions, more than actual genetic paternity.

(4) *Additional References*: We thank the reviewers for suggesting these additional citations and have added them to the manuscript, and reference to paternal care as mating effort (Line 117-118).

(5) *Which Investment Outcomes?*: We have added some text as to why we chose the outcomes we did (admittedly this was somewhat opportunistic as we were using measures we had available from our longitudinal dataset and then constructed the vignette study to map onto those existing measures). We also note as a limitation that there were other measures not included here (e.g. the ones R2 mentions like age at menarche, and others like paternal affect which R2 mentions in the last paragraph of the review) which we did not include. As we were limited in word count we could not include full explanations for all of this, so we offer some additional thoughts here. For age at menarche, while we do have year of menarche we have not been able to get particularly accurate data on the month (as Himba measure time seasonally) and since effect sizes for effects of father absence on menarche are small, we didn't think we would have the resolution to pick them up if they existed. In addition, since omoka children often still live with their fathers, but are sometimes fostered out, links between paternal care and age at menarche become so complex that we would have needed a much larger sample size to detect them.

(6) *ASR:* Himba in Omuhonga have a strongly female-biased ASR, which we do think contributes to the high rate of female concurrent partnerships, divorce and EPP in this population. We are starting to collect more data on spousal separation (which varies by season) and other ecological and social factors which likely contribute to a high rate of EPP being stable in this population. However, we can only nod to this here as this data is still being collected. We added a sentence in the study population section on this. The reviewer also raises the issue of marital type on outcomes in this paper. We don't include marriage type (arranged vs love match) for several reasons: 1) we don't have detailed data on marriage types for many individuals in this study,

particularly adults, 2) while we expect marriage type to impact the frequency of non-paternity within marriage, we don't have any prior expectation of bias father investment as a product of marriage type alone, but that marriage type mediates paternity uncertainty, which we examine as mediating paternal care. We are currently collecting detailed relationship histories for all current partners (marital and non-marital) including details like frequency of sex with non-marital partners. This data is still in early stages, but may help us understand motivations for contributing to and investing in spousal relationships in future papers.

(7) *Livestock Gifts*: We have clarified that the measure we used for Livestock Gifts was a lifetime measure. Although gifts tend to come at set points in an individual's life, we did run the analysis separately for children and adults as children were more likely to have not yet received a gift.

(8) *Brideprice*: The reviewer brings up some very good points about potential confounders in the brideprice analysis. We should first note that brideprice among Himba is quite low compared to other pastoralists, typically only 1-2 cows and 1-2 sheep. Therefore, it is almost always one person who pays the brideprice, either the focal father or someone else (MB, FB). We have added some text in the Study Population section stating this. Borgerhoff Mulder's finding that wealthier men may be marrying women of higher reproductive value is difficult to test directly here because child marriage is common (with many of these brides under the age of 5) and child marriages are very often either never consummated and co-residence never occurs (about a third of the time) or where co-residence occurs for only a few months around menarche, followed by divorce (about another third of cases). If we were to include just the age of the bride in the analysis, it could be misleading in that so many of these marriages never result in children.

(9) *Additional Studies*: We agree with R2 that our vignette results could be expanded upon for a future paper on sexual conflict and appreciate the suggestion as we are thinking more and more about how sexual conflict affects both paternal care in a high EPP setting and how it affects marriage decisions in this context of relatively high female reproductive autonomy. In reference to the second point about paternal affect, this is something we have some opportunistic ethnographic data on but nothing systematic yet. We did try to include quotes in the paper to reinforce the idea that social fatherhood is extremely important, and that men's feelings toward omoka children vary and can change over the child's life depending on the child's value to the household. It would be very interesting in future field seasons to try to measure social closeness between men and their omoka and non-omoka children, and compare this to more quantitative measurements of relative value to the hh and productivity. Anecdotally it seems that closeness of non-marital kids is impacted by issues like labor, where father's report closeness to non-marital kids they find highly value for economic reasons, but this has not been systematically examined at this point.

Similarly, we agree with R2 that changes in the importance of child labor are on the near horizon. Schooling has just started to become more common

among this cohort of children, and we are envisioning a project in the near future looking at these changing dynamics and how they change the roles of fathers (both in that their value to the hh changes when they go to school).

Appendix B

Response to Reviewers

Dear Dr. Hutchinson,

We submit for your consideration a revision of RSPB-2019-2165 entitled, "Paternal care and paternity confidence among Himba pastoralists." We thank the reviewers for their second-round of insightful comments, which we carefully considered as we revised our manuscript. We have made several changes to the manuscript to address your comments, and the comments of the reviewers, which we have detailed below. We appreciate the inclusion of a stats reviewer, as this has helped us to address the concerns of Reviewer 1. We have addressed all reviewer comments, and made several changes. These include:

- **Addressing issues comparing paternity assertions to genetic paternity as raised by reviewer 1.**
- **Removed mention of prediction intervals, and moved those details with additional explanation to the supplementary. This was done because we don't report any prediction intervals in the main manuscript, all results and conclusions are drawn using the full posterior distribution, and differences between posterior distributions, as we describe below. Use of prediction intervals has obviously caused confusion, but are not relevant in the conclusions drawn in this manuscript. Main manuscript figures were adjusted to show 95% intervals.**
- **Included plots of raw data in the supplementary, as suggested by reviewer 3.**

We have detailed responses to reviewer comments are below (in bold text), in addition to track changes in the manuscript. As before, all data and code are available online on the OSF repository.

We again thank you and the reviewers for their time spent on this manuscript.
-SP and BS

Associate Editor Board Member
Comments to Author:
Reviewer 1 has some continuing concerns about the statistical analysis. The statistical reviewer has considered these problems and thinks some additional information needs to be added to make the choice of analysis clearer.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).

As the authors decided to answer the comments from the reviewers in a separate file (which is unusual), it is not readily possible to evaluate their answers in direct relation with the points raised. After some copy/paste work, it appears that most points were considered, but some questions were not addressed. For example, the question "Is the proportion of Omoka and non-omoka individuals analyzed here

different from the rate of non-paternity reported (in the unpublished paper) ?" is left unanswered. Some points require further thoughts.

Authors: We apologize for the confusion created by a creating a separate file. This was meant to aid in responding to reviewer concerns, not to obscure them.

We addressed the issue of proportion of individuals compared to the Scelza et al. 2020 (in press at *Science Advances*) paper in our previous response, and included some additional description in the manuscript. A table in the supplementary materials details the proportion *omoka* for each analysis. Proportion varies based on inclusion/exclusion criteria of the sample as described in the methods section. For example, in the anthropometric analysis, we excluded fostered kids, as fosterage has known independent effects on child health in this population (Scelza & Silk 2014). We should note that these numbers are different than the genetic paternity, which is not an overall proportion but an estimate based on a multilevel model correcting for dyad and individual mother effects.. However, the proportions from the samples used here fall within the expected range of variation of the results from that analysis.

A. The claim that "Fathers have been shown to be very accurate in their paternity assertions..." is now quantitatively substantiated: "..correctly assessing paternity 73% of the time" (line 139). Thus 27% of men are not correctly assessing their paternity status. I would not claim that this is "very accurate".

Authors: We changed "very accurate" to "quite accurate." There is very little data linking genetic paternity data to paternity assertions. In fact, ours (Scelza et al. 2020) is the only study we know of that does so at the individual level. However, Kermyt Anderson published a comprehensive review of paternity studies in 2006 entitled, "How well does paternity confidence match actual paternity?" which looks at men in high-confidence vs low-confidence settings. The low paternity confidence data came from paternity clinics where men were being tested because they disputed their paternity. In these cases, men were not the father 30% of the time, meaning that they were WRONG in their paternity assertions 70% of the time. Comparatively, we see Himba men as being very accurate, but we have tempered our language at the reviewer's request.

B. Lines 207-212: "Here we rely on paternity assertions rather than genetic paternity, as it is the beliefs of men about paternity that are most relevant to investment. However, as these data are from the same set of households used in the genetic paternity study, which showed high accuracy in detecting nonpaternity events, we are confident that there is a strong link between paternity assertions and genetic paternity." That belief (paternity assertions) is stronger than genes (genetic paternity) for paternal investment is perhaps true. That does not mean that genes has no effect. Why not test both, as the data is already there ? Instead of being "confident that there is a strong link between paternity assertions and genetic paternity", please show this link (is this the 73% of correct assessment ??), as the data is already there (or cite a reference). Anyway, the fact that 27% of men have a wrong belief about their paternity is not a strong argument to justify to rely only on paternity assertions...

Authors: We thank the reviewer for raising this issue, as it is an important one, and one that is really at the core of this paper. When we stated that "it is the beliefs about paternity that are most relevant to investment" we are not trying to pit the influences of genes and beliefs against one another. We are more simply referring to the fact that genetic paternity data is not available to men when they are making investment decisions. They are making decisions based on their beliefs about their paternity. Therefore, we maintain that paternity assertions are the most appropriate predictor to use in these analyses. Adding in the genetic data at this point would, we believe be of little added value to the paper and poses significant statistical and logistical challenges including limiting our sample size (as the genetic and

investment samples are not completely overlapping), requiring the addition of several co-authors who worked on the genetic project, and a complicated suite of analyses given that the genetic data were collected using a double-blind method that separates paternity results from demographic and other variables.

C. Again with the prediction interval. I maintain that the sentence (still in the revised version) stating that "All coefficient results and posterior predictions show 89% prediction intervals, to avoid confusion with significance tests." is not correct. I agree that any value (89%, 95%) is arbitrary, but the *justification* expressed here to use 89% is not acceptable to me. This is because, just to avoid a confusion, a different value is chosen, with consequences on results, graphical appearance, etc. A confusion could be avoided by another mean than choosing a different value (e.g. more explanations in the text). A different justification should be provided to stick to this value of 89%. Incidentally, a search on Google Scholar with "89% prediction interval" gives 3 results (yes, only three; to check, do not forget the quotes on the request). A similar search with 90% give 2590 results, and a search with "95% prediction interval" provides around 12000 results. Thus the claim (in the answers to reviewers) that the 89% is "common in modern Bayesian statistics" is not very strong. Please justify properly this choice of 89% which seems, indeed, very unusual in Bayesian statistics.

Authors: We have made an additional note and citation about this issue (also please see response to referee 3). Figures in the main manuscript have been adjusted to reflect 95% intervals instead of 89%. We would like to highlight that the main outcomes in this paper do not rely on prediction intervals as evidence, and instead calculate the posterior differences in different domains (as represented in the density distributions on figures 1-4, and noted and described in the manuscript text as Pr[b>/<0]). In fact, we don't report or rely on 89% intervals in the main manuscript at all, it is only used in the coefficient tables in the supplementary. All results utilize the entire posterior distribution, or differences in posterior distributions. Because of this, we have moved all reference of posterior intervals (and description of why we used them, with citation) to the supplementary materials.

With regard to our assertion that 89% intervals are increasingly common, the reviewer's statement that there are only three papers that have been published using these is false (we have published more than three papers using 89% intervals). The issue here is likely that prediction intervals as we refer to them are often abbreviated as PI, or have a multitude of other names referring to the same principle (confidence interval, credible interval, compatibility interval, percentile interval, posterior interval, high density posterior/prediction interval, or just intervals, etc), so a simple google scholar search likely won't yield accurate results. It is, of course, true that 95% intervals are more common, as a carry-over from frequentist statistics (or, in some cases, many who call OLS regression confidence intervals as prediction intervals), but as previously stated, this is changing to include many different intervals. Ideally a paper might state multiple intervals (50%, 89%, 99% for example), but for our purposes this really isn't necessary, since we don't draw conclusions from percentile intervals. We disagree that it is "very unusual in Bayesian statistics," as a widely used and highly cited statistical book and Bayesian statistical package (McElreath's *Rethinking,* now cited in the text) uses 89% intervals by default. These intervals, and other intervals that are not 95%, have also been used in a number of Royal Society journal publications including Proc B.

 Most importantly, use of other intervals does not change any interpretations in this paper.

D. Where are the figure legends ?

Authors: We included figure legends in our upload, but it appears they were lost or

otherwise were not shown in the final PDF. We apologize for the confusion this error might have caused, and have included figure legends in the manuscript text.

Referee: 3

Comments to the Author(s).
Statistical Review
As asked, I will focus mostly on the statistical aspects of the paper.
I think I can clear up the question why 89% was chosen as a prediction interval: this number appears repeatedly in McElreath's 'Rethinking Statistics' textbook on Bayesian statistics, which is widely used in the social sciences. The author there specifically states that this number was chosen because it is a prime number and easily remembered, and also points out that it is useful to avoid 95% because readers otherwise unconsciously interpret results as though they were hypothesis tests. One thing the same author points out is that in the end, any value chosen here is up to the scientist, and that for normally distributed posteriors (as is the case here), using 80% or 89% or 90% or 95% will give nearly identical results. It would be good if the manuscript's authors could provide this reference in the text. Also, while they do not have to justify the number itself, it would be good to justify why they chose a relatively more relaxed prediction interval than is customary.

Authors: We have added a reference to McElreath's textbook in the section describing the prediction intervals (now moved to the supplementary material, see response to reviewer 1). The reviewer is correct that McElreath's statistical approach inspired much of the approach used in this paper, including prediction intervals. We referenced this text in our response to reviewers in the initial revision. Many of my previous publications use McElreath's statistical package, which defaults to 89% intervals, and for continuity I continue to use them here with a different package. Notably, we don't rely on prediction intervals for our interpretation, and draw our conclusions from the full posterior distributions, as described above. Nevertheless, as noted above figures in the manuscript have been adjusted to illustrate 95% intervals instead of 89% intervals.

With regards to the results, I do not think that the choice of the prediction interval will influence them dramatically, given that the main result is anyways a null result (no difference in treatment). From the plots and estimates, it appears that there is solid evidence for the interpretation given, especially considering that anthropological data are fairly messy and the data sets fairly small and are thereby usually accompanied by a relatively large error.
It would help the reader if the graphs would include a representation of the actual data. It seems that the omoka throughout show much greater variation in estimates; it would be interesting to see whether this is the result of some fathers treating their omokas really bad and others really good (bimodal distribution), or whether there is a continuum in how they are treated.

Authors: We have now included plots in the supplementary to show raw data used in the analyses.

The variation in estimates of omoka outcomes is likely due to variation in treatment of omoka children. As we describe, men may have strong incentives to treat all children equally, but anecdotally we have heard of men treating their omoka children unfavorably, and even cases of men even preferring their omoka children for their hard work and labor. These differences likely result in much variation in child care and investment. It is difficult to see a biomodal distribution, given the type and resolution of the data, and visually the distribution seems continuous rather than biomodal. Additionally, the sample size for omoka is smaller than non-omoka, resulting in higher variance in posterior estimates. However, the variance suggested here is compelling, and requires additional research.