THE ROYAL SOCIETY PUBLISHING

PROCEEDINGS B

A weapons-testes trade-off in males is amplified in female traits

Christine W. Miller, Paul N. Joseph, Rebecca M. Kilner and Zachary Emberts

Article citation details
Proc. R. Soc. B **286**: 20190906.
http://dx.doi.org/10.1098/rspb.2019.0906

Review timeline
Original submission: 16 April 2019
1st revised submission: 23 June 2019
2nd revised submission: 1 July 2019
Final acceptance: 12 July 2019

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

Review History

RSPB-2019-0906.R0 (Original submission)

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?
Excellent

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

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

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
This study explores trade-offs between traits involved in pre-copulatory (i.e. weapons) and post-copulatory (i.e. testes) sexual selection. This field, and indeed the concept of life-history trade-offs more generally, has long been of intense interest to evolutionary biologists. In the current study, the authors provide a novel and clever test of the hypothesis that such trade-offs occur due to traits occupying the same functional category (i.e. reproduction). They achieve this by comparing weapons-testes trade-offs in males to those of homologous structures in females. Their study system, the leaf-footed cactus bug Narnia femorata, allows them to experimentally remove a hind leg (which are used as weapons in reproductive competition by males but not females) during larval stages, and examine resulting development of testes and ovaries. They find that the trade-offs between hind legs and testes in males are echoed, and even amplified, between hind legs and ovaries in female. Given female hind legs have no direct role in reproduction, this suggests that the trade-off is not due to functional grouping of traits.

This is a well-designed and well-executed study, and provides a significant advance to the field (where the mechanisms that underlie trade-offs among sexually selected traits are typically unknown). I particularly commend the authors on their thoughtful and considered interpretations of their results. For example, a question that arose for me while reading was whether increased investment in reproductive output following loss of a limb (autotomy) might reflect an adaptive plastic response rather than a release of resources. While autotomy does not appear to affect offspring-to-adult survival in the lab, in could conceivably lead to reduced lifespan in the wild and/or be a cue of predation risk. Therefore, selection may have acted to favour a plastic increase in early reproductive effort for individuals following autotomy. To the authors' credit, they identify this possibility and devote a paragraph in the Discussion to it. This is a great example of a paper that answers many of your questions as you are reading!

I do have some minor queries/comments for the authors (mostly regarding clarification on some of the analyses), which I list below. I anticipate that these queries will not substantially change the interpretation of the findings.

1) L 35-37: the authors state that studies of life-history traits have provided some surprising results, and cite an empirical paper (reference [4]) to support this. Can the authors briefly state, for non-specialist readers, why this paper provided surprising findings? (e.g. it appears that the cited paper reported a positive effect of low-nutrition diets on immune responses).

2) L 37-38: the authors cite Zera et al. (2001) (reference [4]) for the statement that mechanistic relationships among traits remain unclear. Might it also be worth citing a more recent review to this effect, e.g. Simmons et al. (2017) (reference [8]) might be a good candidate?

3) L 150-152: the autotomy treatment was applied by gripping the left hind legs with forceps then brushing the body with a paintbrush until bugs dropped the leg (as would occur when gripped by a predator). The control group received no experimental manipulation. Can the authors please provide justification for why no handling procedural control was required (e.g. brushing without gripping the legs)? For example, a previous paper from the authors (reference [37]) included several controls (handling, mid-limb removal and baseline) and found no differences in trait investment in these groups. But a reader unfamiliar with this work might wonder if the stress of experimental manipulation affected trait investment.

4) L 153-158: Following experimental manipulation, juveniles were reared in variable sized groups (3-4 or 10-11 individuals per group) with other individuals from the same treatment. I have a couple of questions about this. First, was group size incorporated as a covariate in any analyses? It doesn't appear to be mentioned in the Statistical Analyses, but is it conceivable that group size might have affected investment in certain traits? Second, can the authors provide some further justification of why individuals were raised only with others from the same treatment? Could this have affected trait investment – for example, might interactions (fighting etc.) have differed between groups of autotomized and intact individuals, leading to differential investment?

5) Were the GLMs checked for evidence of overdispersion? Following on from the above comment, if there was any overdispersion then group size might be worth including as an additional explanatory variable to see if it improves model deviance.

6) L 201-203: the authors used a cubic transform for pronotal width, stating that this was to bring the measure to the same scale as mass, before natural log transforming this and all other morphological measures. They cite Tomkins and Simmons (2002) (reference [50]) for support of the cubic transform of pronotum width, but I cannot see mention of such a technique in the cited paper. Is cubic transformation of length/size variables typical when using them as predictors of mass variables typical for this sort of study (e.g. in insect condition literature)? If so, can the authors cite additional supporting references to this effect?

7) L 212: I think this is the first use of the acronym GLM for generalized linear models – if so can authors please include full term here.

8) Can the authors please specify the software used to construct their GLMs?

9) L 252 and L 318-319: in the final model, which was used to directly compare reallocation of investment in gonad mass for males and females, the authors included a covariate for oocyte number that was denoted as a missing value for males. The authors do not describe how the missing values for males were treated by the modelling software. For example, standard techniques include deleting the rows for missing values, or 'imputing' (filling in) the missing values, neither of which are appropriate here. I'm not certain (as I'm no expert in such techniques) about the correct way for dealing with missing values that truly represent a 'not

applicable' score, as is the case here for oocyte number of male individuals. From my searching, it appears some statisticians advocate a 'dummy variable adjustment' method for such cases (e.g. see Note 4 in the Back Matter of this book by Paul Allison https://methods.sagepub.com/book/missing-data or Page 5 of these course notes by Richard Williams https://www3.nd.edu/~rwilliam/stats2/l12.pdf), although it appears this method has been criticised too. Can the authors justify their use and handling of missing values? Alternatively, another possibility might be to correct the females' ovary mass score by the number of oocytes before running this model?

10) L 263-264 and Figure S2: when adjusting for body size in the male testes mass model, the authors found a significant interaction between body size and treatment, as visualised in the supplementary figure. In the Figure S2 caption, the authors describe the result as a more pronounced effect of reallocation following autotomy for smaller males. Do the authors have any thoughts on why this might be the case? This might be an interesting small discussion point (either in the main text or supplementary material, depending on space).

11) Tables 1 and 2: can the direction of the effects (positive or negative) of predictors on response variables be specified? The expected direction for some of these seems intuitive (e.g. I imagine body size vs. ovary mass is a positive relationship), but for example it would be good to report the direction of the relationship between limb muscle mass and gonadal mass (see also L 323-324).

12) L 317-321: I'd suggest combining this model description in the Results with the corresponding part of the Methods section (L 250-252) to avoid repetition.

13) L 382: remove hyphen in "Across-species"?

14) L398-399: sentence beginning "For example" essentially repeats the first sentence of the paragraph. Can it be combined with the following sentence? i.e. "For example, Mocsek and Nijhout [18] found that testes ablation in males led to horn growth……etc".

15) L 436: I wonder if the term "negative covariance" (of gonads and legs) is a bit confusing here given the authors also found a positive relationship between leg muscle mass and gonad mass. Perhaps could be replaced with "trade-off in investment" or something similar to avoid confusing readers? I believe the term "negative covariance" is also used elsewhere in the manuscript to describe the trade-off (e.g. Abstract L21-22); the authors may wish to substitute the terms in these other instances too, but I think it is most important for clarity in the Discussion paragraph that I've identified here.

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.
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
This is a very interesting paper showing a tradeoff between gonadal mass and tissue mass in females and males of the hemipteran, Narnia femorata. This insect can autotomize its hind legs, and when it does, the mass that would have gone into making a larger leg in the adult is instead allocated to make larger gonads. The authors show not only that the tradeoff occurs in both sexes, even though males have substantially larger legs than females, but that, surprisingly, the gain in female ovarian tissues is greater than the gain in male testes. The experiments are well explained and the statistics support the conclusions (although I have a problem with how the stats are presented). Overall a valuable contribution to an interesting biological problem.

I have several issues the authors need to address.

(1). What is meant by phenotypic engineering? This is a brand new term that needs a definition Developmental biologists have used ablation experiments for almost two centuries. If the authors mean something else, then please define the distinction. If not, then this neologism is entirely inappropriate. Just say ablation experiments.

(2). Can you provide pictures of an adult male and female, so the reader can judge the difference between their hind legs? Also give the mean & sd of male and female leg mass in the text where the difference is first described.

(3). There is a conceptual discontinuity between the paragraph that ends on line 58 and the one that begins on line 60.

(4). Lines 63-66: "weapon homologue" is an awful term, and it is totally unclear what is meant here (we have to read several pages along before we are enlightened). Why not be explicit and actually say what you are talking about? Hind legs, used for fighting in males but not in females.

(5). What statistical package or software was used?

(6). Why a chi-squared test instead of the more conventional t-test?

(7). Spell out GLM on first use of the acronym.

(8). Please give the structures of the various GLMs where they are mentioned. Or give a table of the different models and refer to that. The specific models used in Tables 1 and 2 need to be mentioned, perhaps as footnotes. Also give the r-squared and intercept values of the models.

(9). Throughout the paper the authors use covariance to mean that two things co-vary. Covariance is generally used as a statistical term, and has a value. If this is what is meant, then please give the value(s). If the term is used colloquially, then use "co-vary" instead.

(10.). Why is a gigantic font used in lines 273 and 333 ? And what function does the phrase "experimental treatment" serve that could not be put in the figure legend?

(11). Why use log transforms for the data in Figures 2 and 3? It would be much-much clearer if natural numbers are used,

(12). In Table 3 the y-axis is said to represent measures "adjusting for both body size and hind limb muscle mass", but that is not what the axis legend says (it states a plain measure, not a normalized one). Exactly how were these measures "adjusted"? And how are they different from the same measures in Figure 2, where no "adjustment" is mentioned.

(13). Overall the structure of the data in Figures 2 and 3 requires clarification. The caption of Figure 2 is particularly muddled.

(14). It seems to me that the fact that females have a greater response could simply be due to the fact that ovaries are capable of a greater response, by making more large eggs, than testes. Does that seem right? Also, since males have a lesser response, but lose bigger legs, please point out and speculate on where the missing mass in males might go?

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?
Excellent

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

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

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
This study experimentally investigates a potential trade-off between weapons/weapon analogues and gonad (testes/ovary) investment in an insect model, the leaf-footed cactus bug. Specifically, the authors use phenotypic engineering (i.e. removal of one hind leg), which in this species mimics a natural phenomenon, to test the idea that restricting investment in one trait results in increased investment in another. In males, these traits reflect sexually selected traits that function in pre-copulatory (hind leg size) and post-copulatory (gonad/testis size) sexual selection. Such traits are predicted to trade-off according to classic models of sperm competition. In females, the hind leg is a weapon analogue and gonad (ovary) size reflects investment in reproduction. As such, the hind leg is not a trait that functions in reproduction, but instead functions in locomotion. The authors find that removal of the hind leg significantly increases investment in gonads in both males (testis size) and females (ovary mass and oocyte count), and that this effect was also seen when controlling for body size. Moreover, the authors found that the gain in gonadal mass was in fact greater in females than males. The take home message of this paper is that negative covariance between traits in males can extend to homologous traits in females, which suggests that factors other than functional groupings (e.g. reproductive traits) need to be considered in investigations of resource allocation. The authors then posit that factors such as expense of tissues and developmental timing may play a role determining which traits trade-off in this system. Finally, the authors draw attention to the fact that insects with larger legs had larger gonads on average, which raises an interesting discussion of resource acquisition-allocation among individuals.

This paper was an absolute pleasure to read and I think it presents a very important take home message; that is, to consider factors other than functional grouping (i.e. proximity of traits, expense of tissues, developmental timing) in studies of resource allocation. The manuscript is well written, has no major flaws that I can see, and the conclusions are well supported by the

data. As someone who has studied trade-offs between pre- and post-copulatory traits in males, I believe this work adds significantly to our knowledge of resource trade-offs and highlights the need to take a wider perspective on factors that influence trade-offs between traits. It was really enjoyable and insightful to see this perspective presented in the paper and supported by the data. Thus, I think the manuscript will make a valuable contribution to the published literature.

I don't have any major concerns with the manuscript, but do have a number of comments and minor queries. I have identified these by line number below to facilitate a revision of the manuscript.

Line 44: It is not my understanding that the lower branches are prioritized over higher branches in these Y models (admittedly, however, some of the models in de Jong 1993 are quite complicated and perhaps I missed something). Positive and negative covariances are predicted at different branches depending upon when the first major allocation node occurs and whether the traits of interest are above or below this node. Perhaps the authors could clarify this statement or expand upon the statement to make it easier for the reader to follow.

Line 55: On first reading I found myself wanting to know more information about some of these factors, perhaps especially tissue proximity as this idea is less well represented in the literature. Upon reading the discussion, however, I realized all the information I wanted was presented there. I wonder if the authors might be able to find a little more information they could add to the introduction that would satisfy the reader's initial curiosity while, at the same time, not presenting redundant information.

Line 56: citations should be 15-20, not 15, 16-20.

Line 61: I agree with the authors that work on weapon and testes investment has focused on males, but that kind of seems to be a necessity of looking at testes. As such, I find this sentence a little strange. Could the authors broaden this statement? For example, would it also be correct to say that studies of trade-offs in pre- and post-copulatory sexual traits has focused predominately on males?

Line 65: As someone who works on males, I found myself wanting some examples of female weapon homologues at this point in the text. I think adding an example or two would give the reader an idea of how broadly applicable such traits are.

Figure 1: These are great images, though admittedly a little hard to see when printed in B&W, but generally I have no concerns about them. However, I really wonder why the testes are red. Is it pigmentation? Carotenoids?

Line 116: Could the authors add information on the effects larger testes have on fertilization success, e.g. do they sire more offspring? Fertilize more eggs under non-competitive or competitive situations?

Line 120: Please provide information on what traits have been investigated.

Line 129: I find this sentence a little unclear. '…suggesting that trade-offs that span functional groupings of traits.' It is not perfectly clear, in my opinion, whether the authors mean trade-offs occur across multiple traits within a functional grouping (e.g. pre- and post-copulatory traits) or if they mean trade-offs can occur between traits that are considered to represent different functional groupings of traits (e.g. reproduction and locomotion). I think they mean the latter, but it would be nice to see this sentence written a little more clearly.

Line 129-132: I really like this statement and idea presented, and think it is a really important take home point that needs to be added to the literature.

Line 150-152: Does this simulate a predation attempt?

Line 159: I wonder if the fact that individuals were kept in variable sized groups from fourth instar to adulthood may be of relevance to the finding that insects with larger legs had larger gonads. Specifically, while groups were provided with access to the same amount of food, it would have been shared between a different number of individuals, and thus it is possible that resource availability was highly variable between individuals (individuals in groups of 10-11 would likely receive less resources relative to an individual in a group of 3-4). If gonadal investment occurs at some point while individuals are in group housing, this would result in a variable amount of food as well as the suggested (line 441) variation in food quality. Given that variation in resource availability is predicted to turn negative trait covariance to positive covariance, this may go some way towards explain the patterns observed. If there is a reasonable chance that gonadal investment can occur during the period of group housing, I think the authors might consider statistically exploring the effect of group size on their results. This would be in line with numerous studies investigating trait investment under different social conditions that reflect a high chance of sperm competition in the future versus a low chance of future sperm competition.

Line 202: Do you mean cube root transformed?

Results section: I very much like the informative subheading used in the results. Furthermore, setting out the sections clearly by question asked makes the results very easy to follow. I suggest the authors could try and more clearly connect the three questions in the results section to the relevant statistical analysis in the methods section.

Figure 2: Can the authors clarify the symbols used. Are they simply open circles with a standard error? It seems a bit odd that the standard error bars are within the circle in panel (a). I think using a different symbol and presenting mean and standard deviation might make the figure clearer.

Discussion: I really enjoyed the treatment of other factors (e.g. proximity of traits, tissue expense, and developmental timing) in the discussion.

Line 351: Can the authors please clarify the contexts under which fertilization success is boosted? For example, is it under competitive or non-competitive mating scenarios?

Line 390: for the text (out of two), do you mean one hind leg out of two? I assume this is the case, but it isn't entirely clear from the current wording.

Line 439-444: I see the authors have cited a number of relevant papers on resource trade-offs, but I do feel like the following citation is missing from this text.

van Noordwijk, A. J., & de Jong, G. (1986). Acquistion and allocation of resources: their influence on variation in life history tactics. The American Naturalist, 128(1), 137–142.

I think this paper really helps to clearly set out how resource variation can shift negative trait covariances to positive and would be a good addition to this text.

It may also be of interest to the authors (though I am not asking them to include this citation) to

know of a recent paper in which the concept of resource acquisition and allocation where used to extend Parker's models of sperm competition investigating trade-offs between pre- and post-copulatory traits to include variation in resource availability. This work showed that even a relatively small amount of variation in resource availability shifts covariance from negative to positive. Though this is applied to across species patterns, I think it is of some relevance to the positive correlation between pre and post-copulatory sexual traits observed in the current study.

Supriya, K., Price, T. D., & Rowe, M. (2018). Resource variation generates positive correlations between pre- and postcopulatory sexual traits. Behavioral Ecology, 56, 438–7.

Decision letter (RSPB-2019-0906.R0)

03-Jun-2019

Dear Dr Miller:

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.

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,
Victoria Braithwaite

Professor V A Braithwaite
mailto: proceedingsb@royalsociety.org

=====

Associate Editor
Comments to Author:

This manuscript investigates the trade-off between a sexually selected weapon and gonads. They find, as expected from previous data, a negative relationship between weapon size and testes size, where variation in weapon size was induced by experimental leg autotomy. The manipulative approach is nice, but what really makes this manuscript interesting is the fact that they also investigated the trade-off between leg size and ovary size in females. Somewhat surprisingly, the effect of leg autotomy on gonad size was even larger in females. I think this is a quite novel and interesting result, since pre- and post-copulatory trade-offs are typically only investigated in a single sex.

The reviewers all found the manuscript of interest and were generally happy with the experimental design and analysis. However they felt that the statistics should be presented in more detail, and had other comments for clarifying the manuscript. I am therefore recommending revision.

Comments from my own reading of the manuscript, some of which may overlap with comments from the reviewers:
- I would have liked to see some visualization of the relationship between muscle mass lost and gonad size, showing the difference in slope between the sexes.
- In figure 2 the difference between the males (a) looks much smaller than for the females (b), but according to the text these are 17% and 20% differences respectively. This incongruity needs explanation.
- Obviously production of larger legs is costly for females, so it would make sense to discuss this constraint in terms of sexual antagonism, which is currently not mentioned at all.
- They question the simplistic Y-model of allocation, which is good. What about the opportunity for selection on pre- and post-copulatory traits? A male with large weapons might not need large testes if he can monopolize females and avoid sperm competition. A negative covariance in this case might not be a result of a direct energetic trade-off per se, but rather a matter of tailoring investment to the expected intensity of different types of sexual selection.

====

Reviewers' Comments to Author:

Referee: 1

This study explores trade-offs between traits involved in pre-copulatory (i.e. weapons) and post-copulatory (i.e. testes) sexual selection. This field, and indeed the concept of life-history trade-offs more generally, has long been of intense interest to evolutionary biologists. In the current study, the authors provide a novel and clever test of the hypothesis that such trade-offs occur due to traits occupying the same functional category (i.e. reproduction). They achieve this by comparing weapons-testes trade-offs in males to those of homologous structures in females. Their study system, the leaf-footed cactus bug Narnia femorata, allows them to experimentally remove a hind leg (which are used as weapons in reproductive competition by males but not females) during larval stages, and examine resulting development of testes and ovaries. They find that the trade-offs between hind legs and testes in males are echoed, and even amplified, between hind legs and

ovaries in female. Given female hind legs have no direct role in reproduction, this suggests that the trade-off is not due to functional grouping of traits.

This is a well-designed and well-executed study, and provides a significant advance to the field (where the mechanisms that underlie trade-offs among sexually selected traits are typically unknown). I particularly commend the authors on their thoughtful and considered interpretations of their results. For example, a question that arose for me while reading was whether increased investment in reproductive output following loss of a limb (autotomy) might reflect an adaptive plastic response rather than a release of resources. While autotomy does not appear to affect offspring-to-adult survival in the lab, in could conceivably lead to reduced lifespan in the wild and/or be a cue of predation risk. Therefore, selection may have acted to favour a plastic increase in early reproductive effort for individuals following autotomy. To the authors' credit, they identify this possibility and devote a paragraph in the Discussion to it. This is a great example of a paper that answers many of your questions as you are reading!

I do have some minor queries/comments for the authors (mostly regarding clarification on some of the analyses), which I list below. I anticipate that these queries will not substantially change the interpretation of the findings.

1) L 35-37: the authors state that studies of life-history traits have provided some surprising results, and cite an empirical paper (reference [4]) to support this. Can the authors briefly state, for non-specialist readers, why this paper provided surprising findings? (e.g. it appears that the cited paper reported a positive effect of low-nutrition diets on immune responses).

2) L 37-38: the authors cite Zera et al. (2001) (reference [4]) for the statement that mechanistic relationships among traits remain unclear. Might it also be worth citing a more recent review to this effect, e.g. Simmons et al. (2017) (reference [8]) might be a good candidate?

3) L 150-152: the autotomy treatment was applied by gripping the left hind legs with forceps then brushing the body with a paintbrush until bugs dropped the leg (as would occur when gripped by a predator). The control group received no experimental manipulation. Can the authors please provide justification for why no handling procedural control was required (e.g. brushing without gripping the legs)? For example, a previous paper from the authors (reference [37]) included several controls (handling, mid-limb removal and baseline) and found no differences in trait investment in these groups. But a reader unfamiliar with this work might wonder if the stress of experimental manipulation affected trait investment.

4) L 153-158: Following experimental manipulation, juveniles were reared in variable sized groups (3-4 or 10-11 individuals per group) with other individuals from the same treatment. I have a couple of questions about this. First, was group size incorporated as a covariate in any analyses? It doesn't appear to be mentioned in the Statistical Analyses, but is it conceivable that group size might have affected investment in certain traits? Second, can the authors provide some further justification of why individuals were raised only with others from the same treatment? Could this have affected trait investment – for example, might interactions (fighting etc.) have differed between groups of autotomized and intact individuals, leading to differential investment?

5) Were the GLMs checked for evidence of overdispersion? Following on from the above comment, if there was any overdispersion then group size might be worth including as an additional explanatory variable to see if it improves model deviance.

6) L 201-203: the authors used a cubic transform for pronotal width, stating that this was to bring the measure to the same scale as mass, before natural log transforming this and all other

morphological measures. They cite Tomkins and Simmons (2002) (reference [50]) for support of the cubic transform of pronotum width, but I cannot see mention of such a technique in the cited paper. Is cubic transformation of length/size variables typical when using them as predictors of mass variables typical for this sort of study (e.g. in insect condition literature)? If so, can the authors cite additional supporting references to this effect?

7) L 212: I think this is the first use of the acronym GLM for generalized linear models – if so can authors please include full term here.

8) Can the authors please specify the software used to construct their GLMs?

9) L 252 and L 318-319: in the final model, which was used to directly compare reallocation of investment in gonad mass for males and females, the authors included a covariate for oocyte number that was denoted as a missing value for males. The authors do not describe how the missing values for males were treated by the modelling software. For example, standard techniques include deleting the rows for missing values, or 'imputing' (filling in) the missing values, neither of which are appropriate here. I'm not certain (as I'm no expert in such techniques) about the correct way for dealing with missing values that truly represent a 'not applicable' score, as is the case here for oocyte number of male individuals. From my searching, it appears some statisticians advocate a 'dummy variable adjustment' method for such cases (e.g. see Note 4 in the Back Matter of this book by Paul Allison https://methods.sagepub.com/book/missing-data or Page 5 of these course notes by Richard Williams https://www3.nd.edu/~rwilliam/stats2/l12.pdf), although it appears this method has been criticised too. Can the authors justify their use and handling of missing values? Alternatively, another possibility might be to correct the females' ovary mass score by the number of oocytes before running this model?

10) L 263-264 and Figure S2: when adjusting for body size in the male testes mass model, the authors found a significant interaction between body size and treatment, as visualised in the supplementary figure. In the Figure S2 caption, the authors describe the result as a more pronounced effect of reallocation following autotomy for smaller males. Do the authors have any thoughts on why this might be the case? This might be an interesting small discussion point (either in the main text or supplementary material, depending on space).

11) Tables 1 and 2: can the direction of the effects (positive or negative) of predictors on response variables be specified? The expected direction for some of these seems intuitive (e.g. I imagine body size vs. ovary mass is a positive relationship), but for example it would be good to report the direction of the relationship between limb muscle mass and gonadal mass (see also L 323-324).

12) L 317-321: I'd suggest combining this model description in the Results with the corresponding part of the Methods section (L 250-252) to avoid repetition.

13) L 382: remove hyphen in "Across-species"?

14) L398-399: sentence beginning "For example" essentially repeats the first sentence of the paragraph. Can it be combined with the following sentence? i.e. "For example, Mocsek and Nijhout [18] found that testes ablation in males led to horn growth……etc".

15) L 436: I wonder if the term "negative covariance" (of gonads and legs) is a bit confusing here given the authors also found a positive relationship between leg muscle mass and gonad mass. Perhaps could be replaced with "trade-off in investment" or something similar to avoid confusing readers? I believe the term "negative covariance" is also used elsewhere in the

manuscript to describe the trade-off (e.g. Abstract L21-22); the authors may wish to substitute the terms in these other instances too, but I think it is most important for clarity in the Discussion paragraph that I've identified here.

==
Referee: 2

This is a very interesting paper showing a tradeoff between gonadal mass and tissue mass in females and males of the hemipteran, Narnia femorata. This insect can autotomize its hind legs, and when it does, the mass that would have gone into making a larger leg in the adult is instead allocated to make larger gonads. The authors show not only that the tradeoff occurs in both sexes, even though males have substantially larger legs than females, but that, surprisingly, the gain in female ovarian tissues is greater than the gain in male testes. The experiments are well explained and the statistics support the conclusions (although I have a problem with how the stats are presented). Overall a valuable contribution to an interesting biological problem.

I have several issues the authors need to address.

(1). What is meant by phenotypic engineering? This is a brand new term that needs a definition Developmental biologists have used ablation experiments for almost two centuries. If the authors mean something else, then please define the distinction. If not, then this neologism is entirely inappropriate. Just say ablation experiments.

(2). Can you provide pictures of an adult male and female, so the reader can judge the difference between their hind legs? Also give the mean & sd of male and female leg mass in the text where the difference is first described.

(3). There is a conceptual discontinuity between the paragraph that ends on line 58 and the one that begins on line 60.

(4). Lines 63-66: "weapon homologue" is an awful term, and it is totally unclear what is meant here (we have to read several pages along before we are enlightened). Why not be explicit and actually say what you are talking about? Hind legs, used for fighting in males but not in females.

(5). What statistical package or software was used?

(6). Why a chi-squared test instead of the more conventional t-test?

(7). Spell out GLM on first use of the acronym.

(8). Please give the structures of the various GLMs where they are mentioned. Or give a table of the different models and refer to that. The specific models used in Tables 1 and 2 need to be mentioned, perhaps as footnotes. Also give the r-squared and intercept values of the models.

(9). Throughout the paper the authors use covariance to mean that two things co-vary. Covariance is generally used as a statistical term, and has a value. If this is what is meant, then please give the value(s). If the term is used colloquially, then use "co-vary" instead.

(10.). Why is a gigantic font used in lines 273 and 333 ? And what function does the phrase "experimental treatment" serve that could not be put in the figure legend?

(11). Why use log transforms for the data in Figures 2 and 3? It would be much-much clearer if natural numbers are used,

(12). In Table 3 the y-axis is said to represent measures "adjusting for both body size and hind limb muscle mass", but that is not what the axis legend says (it states a plain measure, not a normalized one). Exactly how were these measures "adjusted"? And how are they different from the same measures in Figure 2, where no "adjustment" is mentioned.

(13). Overall the structure of the data in Figures 2 and 3 requires clarification. The caption of Figure 2 is particularly muddled.

(14). It seems to me that the fact that females have a greater response could simply be due to the fact that ovaries are capable of a greater response, by making more large eggs, than testes. Does that seem right? Also, since males have a lesser response, but lose bigger legs, please point out and speculate on where the missing mass in males might go?

==

Referee: 3

This study experimentally investigates a potential trade-off between weapons/weapon analogues and gonad (testes/ovary) investment in an insect model, the leaf-footed cactus bug. Specifically, the authors use phenotypic engineering (i.e. removal of one hind leg), which in this species mimics a natural phenomenon, to test the idea that restricting investment in one trait results in increased investment in another. In males, these traits reflect sexually selected traits that function in pre-copulatory (hind leg size) and post-copulatory (gonad/testis size) sexual selection. Such traits are predicted to trade-off according to classic models of sperm competition. In females, the hind leg is a weapon analogue and gonad (ovary) size reflects investment in reproduction. As such, the hind leg is not a trait that functions in reproduction, but instead functions in locomotion. The authors find that removal of the hind leg significantly increases investment in gonads in both males (testis size) and females (ovary mass and oocyte count), and that this effect was also seen when controlling for body size. Moreover, the authors found that the gain in gonadal mass was in fact greater in females than males. The take home message of this paper is that negative covariance between traits in males can extend to homologous traits in females, which suggests that factors other than functional groupings (e.g. reproductive traits) need to be considered in investigations of resource allocation. The authors then posit that factors such as expense of tissues and developmental timing may play a role determining which traits trade-off in this system. Finally, the authors draw attention to the fact that insects with larger legs had larger gonads on average, which raises an interesting discussion of resource acquisition-allocation among individuals.

This paper was an absolute pleasure to read and I think it presents a very important take home message; that is, to consider factors other than functional grouping (i.e. proximity of traits, expense of tissues, developmental timing) in studies of resource allocation. The manuscript is well written, has no major flaws that I can see, and the conclusions are well supported by the data. As someone who has studied trade-offs between pre- and post-copulatory traits in males, I believe this work adds significantly to our knowledge of resource trade-offs and highlights the need to take a wider perspective on factors that influence trade-offs between traits. It was really enjoyable and insightful to see this perspective presented in the paper and supported by the data. Thus, I think the manuscript will make a valuable contribution to the published literature.

I don't have any major concerns with the manuscript, but do have a number of comments and minor queries. I have identified these by line number below to facilitate a revision of the manuscript.

Line 44: It is not my understanding that the lower branches are prioritized over higher branches in these Y models (admittedly, however, some of the models in de Jong 1993 are quite complicated and perhaps I missed something). Positive and negative covariances are predicted at different branches depending upon when the first major allocation node occurs and whether the traits of interest are above or below this node. Perhaps the authors could clarify this statement or expand upon the statement to make it easier for the reader to follow.

Line 55: On first reading I found myself wanting to know more information about some of these factors, perhaps especially tissue proximity as this idea is less well represented in the literature. Upon reading the discussion, however, I realized all the information I wanted was presented there. I wonder if the authors might be able to find a little more information they could add to the introduction that would satisfy the reader's initial curiosity while, at the same time, not presenting redundant information.

Line 56: citations should be 15-20, not 15, 16-20.

Line 61: I agree with the authors that work on weapon and testes investment has focused on males, but that kind of seems to be a necessity of looking at testes. As such, I find this sentence a little strange. Could the authors broaden this statement? For example, would it also be correct to say that studies of trade-offs in pre- and post-copulatory sexual traits has focused predominately on males?

Line 65: As someone who works on males, I found myself wanting some examples of female weapon homologues at this point in the text. I think adding an example or two would give the reader an idea of how broadly applicable such traits are.

Figure 1: These are great images, though admittedly a little hard to see when printed in B&W, but generally I have no concerns about them. However, I really wonder why the testes are red. Is it pigmentation? Carotenoids?

Line 116: Could the authors add information on the effects larger testes have on fertilization success, e.g. do they sire more offspring? Fertilize more eggs under non-competitive or competitive situations?

Line 120: Please provide information on what traits have been investigated.

Line 129: I find this sentence a little unclear. '…suggesting that trade-offs that span functional groupings of traits.' It is not perfectly clear, in my opinion, whether the authors mean trade-offs occur across multiple traits within a functional grouping (e.g. pre- and post-copulatory traits) or if they mean trade-offs can occur between traits that are considered to represent different functional groupings of traits (e.g. reproduction and locomotion). I think they mean the latter, but it would be nice to see this sentence written a little more clearly.

Line 129-132: I really like this statement and idea presented, and think it is a really important take home point that needs to be added to the literature.

Line 150-152: Does this simulate a predation attempt?

Line 159: I wonder if the fact that individuals were kept in variable sized groups from fourth instar to adulthood may be of relevance to the finding that insects with larger legs had larger gonads. Specifically, while groups were provided with access to the same amount of food, it would have been shared between a different number of individuals, and thus it is possible that resource availability was highly variable between individuals (individuals in groups of 10-11 would likely receive less resources relative to an individual in a group of 3-4). If gonadal investment occurs at some point while individuals are in group housing, this would result in a variable amount of food as well as the suggested (line 441) variation in food quality. Given that variation in resource availability is predicted to turn negative trait covariance to positive covariance, this may go some way towards explain the patterns observed. If there is a reasonable chance that gonadal investment can occur during the period of group housing, I think the authors might consider statistically exploring the effect of group size on their results. This would be in line with numerous studies investigating trait investment under different social conditions that reflect a high chance of sperm competition in the future versus a low chance of future sperm competition.

Line 202: Do you mean cube root transformed?

Results section: I very much like the informative subheading used in the results. Furthermore, setting out the sections clearly by question asked makes the results very easy to follow. I suggest the authors could try and more clearly connect the three questions in the results section to the relevant statistical analysis in the methods section.

Figure 2: Can the authors clarify the symbols used. Are they simply open circles with a standard error? It seems a bit odd that the standard error bars are within the circle in panel (a). I think using a different symbol and presenting mean and standard deviation might make the figure clearer.

Discussion: I really enjoyed the treatment of other factors (e.g. proximity of traits, tissue expense, and developmental timing) in the discussion.

Line 351: Can the authors please clarify the contexts under which fertilization success is boosted? For example, is it under competitive or non-competitive mating scenarios?

Line 390: for the text (out of two), do you mean one hind leg out of two? I assume this is the case, but it isn't entirely clear from the current wording.

Line 439-444: I see the authors have cited a number of relevant papers on resource trade-offs, but I do feel like the following citation is missing from this text.

van Noordwijk, A. J., & de Jong, G. (1986). Acquistion and allocation of resources: their influence on variation in life history tactics. The American Naturalist, 128(1), 137–142.

I think this paper really helps to clearly set out how resource variation can shift negative trait covariances to positive and would be a good addition to this text.

It may also be of interest to the authors (though I am not asking them to include this citation) to know of a recent paper in which the concept of resource acquisition and allocation where used to extend Parker's models of sperm competition investigating trade-offs between pre- and post-copulatory traits to include variation in resource availability. This work showed that even a relatively small amount of variation in resource availability shifts covariance from negative to

positive. Though this is applied to across species patterns, I think it is of some relevance to the positive correlation between pre and post-copulatory sexual traits observed in the current study.

Supriya, K., Price, T. D., & Rowe, M. (2018). Resource variation generates positive correlations between pre- and postcopulatory sexual traits. Behavioral Ecology, 56, 438–7.

Author's Response to Decision Letter for (RSPB-2019-0906.R0)

See Appendix A.

Decision letter (RSPB-2019-0906.R1)

28-Jun-2019

Dear Dr Miller

We are pleased to inform you that your manuscript RSPB-2019-0906.R1 entitled "A weapons-testes trade-off in males is amplified in female traits" has been accepted for publication in Proceedings B.

The Associate Editor has recommended publication, but also suggests some minor revisions to your manuscript. Therefore, we invite you to respond to the 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,

Proceedings B

==========
mailto:proceedingsb@royalsociety.org

==========
Associate Editor:

Comments to Author:

The authors have done a thorough job addressing all comments from the previous round of review. I just have a quibble with the new paragraph in the discussion which mentions sexual antagonism. On lines 431-433 it's stated that sexual antagonism can't explain the additional ovarian mass females gain after loss of a hind limb. I'm not sure of the reasoning here. To me, this is good evidence that production of large hind limbs is not advantageous in females, and simply occurs as a developmental constraint because they are strongly selected in males. This would then be an example of sexual antagonism (or ontogenetic/intralocus sexual conflict, if you will) over limb size. If you disagree, please clarify why, and if you agree, please amend this statement.

Decision letter (RSPB-2019-0906.R2)

12-Jul-2019

Dear Dr Miller

I am pleased to inform you that your manuscript entitled "A weapons-testes trade-off in males is amplified in female traits" 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

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,
Victoria Braithwaite

==========

Professor V A Braithwaite
Editor, Proceedings B
mailto: proceedingsb@royalsociety.org

Appendix A

Dear Associate Editor and Reviewers:

Thank you for your helpful and thorough comments on our manuscript. We are very happy that you found this manuscript to be valuable, and we appreciate the positive comments. Below we provide referee comments, our responses, and we provide line numbers to direct the referees to changes in the revised document. In addressing the comments, we found our document exceeding the page limit. Thus, we have also made some changes to reduce the length of the article, including cutting our original Table 1 and Figure 2 and moving these results to the text. A "track-changes" version of our article is appended at the end of this document so that the changes between the original and current version can be examined.

Associate Editor comments:

- I would have liked to see some visualization of the relationship between muscle mass lost and gonad size, showing the difference in slope between the sexes.

Thank you for this insight; it has resulted in a new panel in Figure 2 and an expanded test of interactions, now shown in Table 1.

- In figure 2 the difference between the males (a) looks much smaller than for the females (b), but according to the text these are 17% and 20% differences respectively. This incongruity needs explanation.

We now explain in the text in more detail how we calculated our estimates of relative gonadal gain following autotomy (Lines 305-319).

- Obviously production of larger legs is costly for females, so it would make sense to discuss this constraint in terms of sexual antagonism, which is currently not mentioned at all.

Lines 428-439. We added a new paragraph to the Discussion.

- They question the simplistic Y-model of allocation, which is good. What about the opportunity for selection on pre- and post-copulatory traits? A male with large weapons might not need large testes if he can monopolize females and avoid sperm competition. A negative covariance in this case might not be a result of a direct energetic trade-off per se, but rather a matter of tailoring investment to the expected intensity of different types of sexual selection.

Lines 435-437. We now discuss allocation priorities and how they likely differ for each sex.

Referee 1:

This study explores trade-offs between traits involved in pre-copulatory (i.e. weapons) and post-copulatory (i.e. testes) sexual selection. This field, and indeed the concept of life-history trade-offs more generally, has long been of intense interest to evolutionary biologists. In the current study, the authors provide a novel and clever test of the hypothesis that such trade-offs occur due to traits occupying the same functional category (i.e. reproduction). They achieve this by comparing weapons-testes trade-offs in males to those of homologous structures in females.

Their study system, the leaf-footed cactus bug *Narnia femorata*, allows them to experimentally remove a hind leg (which are used as weapons in reproductive competition by males but not females) during larval stages, and examine resulting development of testes and ovaries. They find that the trade-offs between hind legs and testes in males are echoed, and even amplified, between hind legs and ovaries in female. Given female hind legs have no direct role in reproduction, this suggests that the trade-off is not due to functional grouping of traits.

This is a well-designed and well-executed study, and provides a significant advance to the field (where the mechanisms that underlie trade-offs among sexually selected traits are typically unknown). I particularly commend the authors on their thoughtful and considered interpretations of their results. For example, a question that arose for me while reading was whether increased investment in reproductive output following loss of a limb (autotomy) might reflect an adaptive plastic response rather than a release of resources. While autotomy does not appear to affect offspring-to-adult survival in the lab, in could conceivably lead to reduced lifespan in the wild and/or be a cue of predation risk. Therefore, selection may have acted to favour a plastic increase in early reproductive effort for individuals following autotomy. To the authors' credit, they identify this possibility and devote a paragraph in the Discussion to it. This is a great example of a paper that answers many of your questions as you are reading!

I do have some minor queries/comments for the authors (mostly regarding clarification on some of the analyses), which I list below. I anticipate that these queries will not substantially change the interpretation of the findings.

1) L 35-37: the authors state that studies of life-history traits have provided some surprising results, and cite an empirical paper (reference [4]) to support this. Can the authors briefly state, for non-specialist readers, why this paper provided surprising findings? (e.g. it appears that the cited paper reported a positive effect of low-nutrition diets on immune responses).

Line 36. We realized one or a handful of citations would not due this statement justice, and so we now cite review/idea papers that discuss this issue.

2) L 37-38: the authors cite Zera et al. (2001) (reference [4]) for the statement that mechanistic relationships among traits remain unclear. Might it also be worth citing a more recent review to this effect, e.g. Simmons et al. (2017) (reference [8]) might be a good candidate?

Line 37. We have added this reference.

3) L 150-152: the autotomy treatment was applied by gripping the left hind legs with forceps then brushing the body with a paintbrush until bugs dropped the leg (as would occur when gripped by a predator). The control group received no experimental manipulation. Can the authors please provide justification for why no handling procedural control was required (e.g.

brushing without gripping the legs)? For example, a previous paper from the authors (reference [37]) included several controls (handling, mid-limb removal and baseline) and found no differences in trait investment in these groups. But a reader unfamiliar with this work might wonder if the stress of experimental manipulation affected trait investment.

Lines 141-145. We have added a brief description of the multiple controls used in the previous study, and why only one control was used here.

4) L 153-158: Following experimental manipulation, juveniles were reared in variable sized groups (3-4 or 10-11 individuals per group) with other individuals from the same treatment. I have a couple of questions about this. First, was group size incorporated as a covariate in any analyses? It doesn't appear to be mentioned in the Statistical Analyses, but is it conceivable that group size might have affected investment in certain traits?

We are very interested in this question, too. We considered including developmental group size in these analyses but realized that this topic deserved its own paper (and would make the current paper too bulky). We are in the process of completing two papers on group size effects using this dataset in combination with other studies.

Not including group size as a covariate in the current analyses adds extra noise to our results, and in this way our results should be more conservative.

Second, can the authors provide some further justification of why individuals were raised only with others from the same treatment? Could this have affected trait investment – for example, might interactions (fighting etc.) have differed between groups of autotomized and intact individuals, leading to differential investment?

This is a fascinating idea to consider. In this case, it was logistically necessary for us to raise individuals with others from the same treatment. We have now explained in the text the logistical constraints that led to this approach (lines 156-161). However, it is entirely possible that social interactions may differ for these groups, and we would like to explore this in the future.

5) Were the GLMs checked for evidence of overdispersion? Following on from the above comment, if there was any overdispersion then group size might be worth including as an additional explanatory variable to see if it improves model deviance.

We initially used the Poisson distribution for our test of the number of eggs produced by females in this study. These data were indeed over-dispersed, and so we have now opted for a negative binomial distribution (Lines 202, 228-229).

6) L 201-203: the authors used a cubic transform for pronotal width, stating that this was to bring the measure to the same scale as mass, before natural log transforming this and all other morphological measures. They cite Tomkins and Simmons (2002) (reference [50]) for support of

the cubic transform of pronotum width, but I cannot see mention of such a technique in the cited paper. Is cubic transformation of length/size variables typical when using them as predictors of mass variables typical for this sort of study (e.g. in insect condition literature)? If so, can the authors cite additional supporting references to this effect?

Thank you for catching this. We obtained the same results when removing the cubic transformation, and so have removed this line and citation.

7) L 212: I think this is the first use of the acronym GLM for generalized linear models – if so can authors please include full term here.

Done, line 208.

8) Can the authors please specify the software used to construct their GLMs?

Done, lines 202-203.

9) L 252 and L 318-319: in the final model, which was used to directly compare reallocation of investment in gonad mass for males and females, the authors included a covariate for oocyte number that was denoted as a missing value for males. The authors do not describe how the missing values for males were treated by the modelling software. For example, standard techniques include deleting the rows for missing values, or 'imputing' (filling in) the missing values, neither of which are appropriate here. I'm not certain (as I'm no expert in such techniques) about the correct way for dealing with missing values that truly represent a 'not applicable' score, as is the case here for oocyte number of male individuals. From my searching, it appears some statisticians advocate a 'dummy variable adjustment' method for such cases (e.g. see Note 4 in the Back Matter of this book by Paul Allison https://urldefense.proofpoint.com/v2/url?u=https-3A__methods.sagepub.com_book_missing-2Ddata&d=DwIFaQ&c=sJ6xIWYx-zLMB3EPkvcnVg&r=RobHzc26Y6FIBGQXrNSdpg&m=597DlWbK7LWhW0WJRgj5GGezjQ-urbdzT7sFX92nnP0&s=fFuvR66VBhY_KVS5hy8YoibpITInNzJOOe8SrOZGczc&e= or Page 5 of these course notes by Richard Williams https://urldefense.proofpoint.com/v2/url?u=https-3A__www3.nd.edu_-7Erwilliam_stats2_l12.pdf&d=DwIFaQ&c=sJ6xIWYx-zLMB3EPkvcnVg&r=RobHzc26Y6FIBGQXrNSdpg&m=597DlWbK7LWhW0WJRgj5GGezjQ-urbdzT7sFX92nnP0&s=Fsghdqk1x_u4ZJOJZ7zdShHhuHPnhsoatpsIk6d8_yg&e=), although it appears this method has been criticised too. Can the authors justify their use and handling of missing values? Alternatively, another possibility might be to correct the females' ovary mass score by the number of oocytes before running this model?

Thank you for your detective work on this matter. It doesn't appear that there is a single and ideal solution for a direct comparison of the sexes in this case, where oocyte count is a required covariate for females but not for males. We gave this some time and decided the best approach would be to ask our question in both a quantitative and qualitative manner. In addition to the full model that directly compares the sexes (left column in Table 1), we have now constructed models unique to each sex (right columns in Table 1) to make fewer assumptions. In all

approaches we see a consistent pattern, where the gain in gonadal mass is greater for females following autotomy. (Figure 2a, Lines 305-310, 313-320)

10) L 263-264 and Figure S2: when adjusting for body size in the male testes mass model, the authors found a significant interaction between body size and treatment, as visualised in the supplementary figure. In the Figure S2 caption, the authors describe the result as a more pronounced effect of reallocation following autotomy for smaller males. Do the authors have any thoughts on why this might be the case? This might be an interesting small discussion point (either in the main text or supplementary material, depending on space).

Because of space considerations, we have now added some thoughts to the supplementary material document – it is an interesting question!

11) Tables 1 and 2: can the direction of the effects (positive or negative) of predictors on response variables be specified? The expected direction for some of these seems intuitive (e.g. I imagine body size vs. ovary mass is a positive relationship), but for example it would be good to report the direction of the relationship between limb muscle mass and gonadal mass (see also L 323-324).

As mentioned above in the response to the Associate Editor, we have now provided a graph of this relationship (figure 2*a*).

12) L 317-321: I'd suggest combining this model description in the Results with the corresponding part of the Methods section (L 250-252) to avoid repetition.

We have attempted to minimize these and other redundancies.

13) L 382: remove hyphen in "Across-species"?

Done.

14) L398-399: sentence beginning "For example" essentially repeats the first sentence of the paragraph. Can it be combined with the following sentence? i.e. "For example, Mocsek and Nijhout [18] found that testes ablation in males led to horn growth……etc".

Done.

15) L 436: I wonder if the term "negative covariance" (of gonads and legs) is a bit confusing here given the authors also found a positive relationship between leg muscle mass and gonad mass. Perhaps could be replaced with "trade-off in investment" or something similar to avoid confusing readers? I believe the term "negative covariance" is also used elsewhere in the manuscript to describe the trade-off (e.g. Abstract L21-22); the authors may wish to substitute the terms in these other instances too, but I think it is most important for clarity in the Discussion paragraph that I've identified here.

Substitutions made in the specified spots, including Lines 21-22, 419, and more.

Referee 2:

This is a very interesting paper showing a tradeoff between gonadal mass and tissue mass in females and males of the hemipteran, Narnia femorata. This insect can autotomize its hind legs, and when it does, the mass that would have gone into making a larger leg in the adult is instead allocated to make larger gonads. The authors show not only that the tradeoff occurs in both sexes, even though males have substantially larger legs than females, but that, surprisingly, the gain in female ovarian tissues is greater than the gain in male testes. The experiments are well explained and the statistics support the conclusions (although I have a problem with how the stats are presented). Overall a valuable contribution to an interesting biological problem.

I have several issues the authors need to address.

(1). What is meant by phenotypic engineering? This is a brand new term that needs a definition Developmental biologists have used ablation experiments for almost two centuries. If the authors mean something else, then please define the distinction. If not, then this neologism is entirely inappropriate. Just say ablation experiments.

Lines 66-68, 98-102. We have now added that our methods resemble established ablation experiments, and we are glad to know that these have been used for almost two centuries (and have added that in the text). In the first mention of the term “phenotypic engineering” we include a citation to Zera and Harshmann (2001) where this procedure is reviewed and the term is defined.

(2). Can you provide pictures of an adult male and female, so the reader can judge the difference between their hind legs? Also give the mean & sd of male and female leg mass in the text where the difference is first described.

We have added a figure (figure S1) and the mean & sd to the electronic supplementary materials.

(3). There is a conceptual discontinuity between the paragraph that ends on line 58 and the one that begins on line 60.

Lines 54-55. This transition has been improved with the addition of a sentence at the end of the paragraph.

(4). Lines 63-66: "weapon homologue" is an awful term, and it is totally unclear what is meant here (we have to read several pages along before we are enlightened). Why not be explicit and actually say what you are talking about? Hind legs, used for fighting in males but not in females.

Lines 56-65. We have attempted to improve the wording here to make it less distasteful, and we also added a short and familiar example to help clarify what we mean (female horns in some African ungulates).

(5). What statistical package or software was used?

Added in Lines 202-203.

(6). Why a chi-squared test instead of the more conventional t-test?

All models run are generalized linear models using a Wald Chi-square statistic.

(7). Spell out GLM on first use of the acronym.

Done, line 217.

(8). Please give the structures of the various GLMs where they are mentioned. Or give a table of the different models and refer to that. The specific models used in Tables 1 and 2 need to be mentioned, perhaps as footnotes. Also give the r-squared and intercept values of the models.

We now clearly state that the models shown in table 1 include the model structure, and in the text we have made sure that additional details are included.

(9). Throughout the paper the authors use covariance to mean that two things co-vary. Covariance is generally used as a statistical term, and has a value. If this is what is meant, then please give the value(s). If the term is used colloquially, then use "co-vary" instead.

“Negative covariance” was changed to “negative correlation” or similar throughout the document (e.g. Lines 71, 76, 111, 396, 417, 424)

(10.). Why is a gigantic font used in lines 273 and 333 ? And what function does the phrase "experimental treatment" serve that could not be put in the figure legend?

We have removed “Experimental treatment” from the x-axis labels in the figure.

(11). Why use log transforms for the data in Figures 2 and 3? It would be much-much clearer if natural numbers are used,

Natural numbers, including back transformations, are now seen in figures 2a, S2, and S3. The others (the new figure 2b,c) directly report model output which we also think is valuable to report. These are estimated marginal means directly from the GLM, adjusting for body size and other factors (Table 1 left column).

(12). In Table 3 the y-axis is said to represent measures "adjusting for both body size and hind limb muscle mass", but that is not what the axis legend says (it states a plain measure, not a normalized one). Exactly how were these measures "adjusted"? And how are they different from the same measures in Figure 2, where no "adjustment" is mentioned.

We have now completely rewritten the Figure caption and our explanations.

(13). Overall the structure of the data in Figures 2 and 3 requires clarification. The caption of Figure 2 is particularly muddled.

We have clarified our captions and made clear links with the models used.

(14). It seems to me that the fact that females have a greater response could simply be due to the fact that ovaries are capable of a greater response, by making more large eggs, than testes. Does

that seem right? Also, since males have a lesser response, but lose bigger legs, please point out and speculate on where the missing mass in males might go?

Lines 429-439. We've now added a new paragraph to the Discussion section to address the related comments of other reviewers and the associate editor.

Referee 3:

This study experimentally investigates a potential trade-off between weapons/weapon analogues and gonad (testes/ovary) investment in an insect model, the leaf-footed cactus bug. Specifically, the authors use phenotypic engineering (i.e. removal of one hind leg), which in this species mimics a natural phenomenon, to test the idea that restricting investment in one trait results in increased investment in another. In males, these traits reflect sexually selected traits that function in pre-copulatory (hind leg size) and post-copulatory (gonad/testis size) sexual selection. Such traits are predicted to trade-off according to classic models of sperm competition. In females, the hind leg is a weapon analogue and gonad (ovary) size reflects investment in reproduction. As such, the hind leg is not a trait that functions in reproduction, but instead functions in locomotion. The authors find that removal of the hind leg significantly increases investment in gonads in both males (testis size) and females (ovary mass and oocyte count), and that this effect was also seen when controlling for body size. Moreover, the authors found that the gain in gonadal mass was in fact greater in females than males. The take home message of this paper is that negative covariance between traits in males can extend to homologous traits in females, which suggests that factors other than functional groupings (e.g. reproductive traits) need to be considered in investigations of resource allocation. The authors then posit that factors such as expense of tissues and developmental timing may play a role determining which traits trade-off in this system. Finally, the authors draw attention to the fact that insects with larger legs had larger gonads on average, which raises an interesting discussion of resource acquisition-allocation among individuals.

This paper was an absolute pleasure to read and I think it presents a very important take home message; that is, to consider factors other than functional grouping (i.e. proximity of traits, expense of tissues, developmental timing) in studies of resource allocation. The manuscript is well written, has no major flaws that I can see, and the conclusions are well supported by the data. As someone who has studied trade-offs between pre- and post-copulatory traits in males, I believe this work adds significantly to our knowledge of resource trade-offs and highlights the need to take a wider perspective on factors that influence trade-offs between traits. It was really enjoyable and insightful to see this perspective presented in the paper and supported by the data. Thus, I think the manuscript will make a valuable contribution to the published literature.

I don't have any major concerns with the manuscript, but do have a number of comments and minor queries. I have identified these by line number below to facilitate a revision of the manuscript.

Line 44: It is not my understanding that the lower branches are prioritized over higher branches in these Y models (admittedly, however, some of the models in de Jong 1993 are quite complicated and perhaps I missed something). Positive and negative covariances are predicted at different branches depending upon when the first major allocation node occurs and whether the traits of interest are above or below this node. Perhaps the authors could clarify this statement or expand upon the statement to make it easier for the reader to follow.

Line 39-40. We have now modified our words in this section.

Line 55: On first reading I found myself wanting to know more information about some of these factors, perhaps especially tissue proximity as this idea is less well represented in the literature. Upon reading the discussion, however, I realized all the information I wanted was presented there. I wonder if the authors might be able to find a little more information they could add to the introduction that would satisfy the reader's initial curiosity while, at the same time, not presenting redundant information.

We agree that it would be great to add more description to the introduction. We tried to do this; unfortunately, we realized that our explanation took us over the word limit, even with trying to cut back in other areas.

Line 56: citations should be 15-20, not 15, 16-20.

Fixed, thank you.

Line 61: I agree with the authors that work on weapon and testes investment has focused on males, but that kind of seems to be a necessity of looking at testes. As such, I find this sentence a little strange. Could the authors broaden this statement? For example, would it also be correct to say that studies of trade-offs in pre- and post-copulatory sexual traits has focused predominately on males?

Lines 57-58. This is a nice improvement! Changed as suggested.

Line 65: As someone who works on males, I found myself wanting some examples of female weapon homologues at this point in the text. I think adding an example or two would give the reader an idea of how broadly applicable such traits are.

Lines 59-62. This is a useful addition.

Figure 1: These are great images, though admittedly a little hard to see when printed in B&W, but generally I have no concerns about them. However, I really wonder why the testes are red. Is it pigmentation? Carotenoids?

We find this fascinating, too. It does appear to be pigmentation, and we think it is probably carotenoids. These testes are found just under the dorsal cuticle, and we wonder if those

populations routinely exposed to bright sunshine have added the pigmentation as a protectant against UV. We have identified some understory-living species that have white testes.

Line 116: Could the authors add information on the effects larger testes have on fertilization success, e.g. do they sire more offspring? Fertilize more eggs under non-competitive or competitive situations?

Lines 109, 341.

Line 120: Please provide information on what traits have been investigated.

Lines 113-114. We have now listed several of these traits.

Line 129: I find this sentence a little unclear. '…suggesting that trade-offs that span functional groupings of traits.' It is not perfectly clear, in my opinion, whether the authors mean trade-offs occur across multiple traits within a functional grouping (e.g. pre- and post-copulatory traits) or if they mean trade-offs can occur between traits that are considered to represent different functional groupings of traits (e.g. reproduction and locomotion). I think they mean the latter, but it would be nice to see this sentence written a little more clearly.

Line 121. This has been rewritten as suggested

Line 129-132: I really like this statement and idea presented, and think it is a really important take home point that needs to be added to the literature.

Line 150-152: Does this simulate a predation attempt?

We have now added explanation in Lines 94-96 that it occurs as a means to escape. In many cases, they use autotomy to escape a bad molt, where they cannot otherwise free themselves from a shed cuticle. Our approach was to minimize stress to the insect as much as possible, so saying it similates a predation attempt might not be exactly correct. It does simulate entrapment of a limb.

Line 159: I wonder if the fact that individuals were kept in variable sized groups from fourth instar to adulthood may be of relevance to the finding that insects with larger legs had larger gonads. Specifically, while groups were provided with access to the same amount of food, it would have been shared between a different number of individuals, and thus it is possible that resource availability was highly variable between individuals (individuals in groups of 10-11 would likely receive less resources relative to an individual in a group of 3-4). If gonadal investment occurs at some point while individuals are in group housing, this would result in a variable amount of food as well as the suggested (line 441) variation in food quality. Given that variation in resource availability is predicted to turn negative trait covariance to positive covariance, this may go some way towards explain the patterns observed. If there is a reasonable

chance that gonadal investment can occur during the period of group housing, I think the authors might consider statistically exploring the effect of group size on their results. This would be in line with numerous studies investigating trait investment under different social conditions that reflect a high chance of sperm competition in the future versus a low chance of future sperm competition.

We initially considered including developmental group size in these analyses but realized that this topic deserved its own paper (and would make the current paper too bulky). We are in the process of completing two papers on group size effects using this dataset in combination with other studies.

Excluding group size as a covariate in the current analyses adds extra noise to our results, and in this way our results should be more conservative.

Line 202: Do you mean cube root transformed?

We dropped this transformation at the suggestion of another reviewer.

Results section: I very much like the informative subheading used in the results. Furthermore, setting out the sections clearly by question asked makes the results very easy to follow. I suggest the authors could try and more clearly connect the three questions in the results section to the relevant statistical analysis in the methods section.

We have now added informative subheadings to the statistical methods section.

Figure 2: Can the authors clarify the symbols used. Are they simply open circles with a standard error? It seems a bit odd that the standard error bars are within the circle in panel (a). I think using a different symbol and presenting mean and standard deviation might make the figure clearer.

It's certainly not every day that standard errors are so small as to be subsumed by the marker. We have now clarified that hollow circles denote the EMM $\pm$ standard error bars.

Discussion: I really enjoyed the treatment of other factors (e.g. proximity of traits, tissue expense, and developmental timing) in the discussion.

Line 351: Can the authors please clarify the contexts under which fertilization success is boosted? For example, is it under competitive or non-competitive mating scenarios?

Line 339: We have added a sentence explaining that fertilization success is boosted for these males in non-competitive mating scenarios.

Line 390: for the text (out of two), do you mean one hind leg out of two? I assume this is the case, but it isn't entirely clear from the current wording.

Line 378. We removed the "(out of two)" to clarify the statement.

Line 439-444: I see the authors have cited a number of relevant papers on resource trade-offs, but I do feel like the following citation is missing from this text.

van Noordwijk, A. J., & de Jong, G. (1986). Acquistion and allocation of resources: their influence on variation in life history tactics. The American Naturalist, 128(1), 137–142.

I think this paper really helps to clearly set out how resource variation can shift negative trait covariances to positive and would be a good addition to this text.

This is an important paper to include, which we have now done.

It may also be of interest to the authors (though I am not asking them to include this citation) to know of a recent paper in which the concept of resource acquisition and allocation where used to extend Parker's models of sperm competition investigating trade-offs between pre- and post-copulatory traits to include variation in resource availability. This work showed that even a relatively small amount of variation in resource availability shifts covariance from negative to positive. Though this is applied to across species patterns, I think it is of some relevance to the positive correlation between pre and post-copulatory sexual traits observed in the current study.

Supriya, K., Price, T. D., & Rowe, M. (2018). Resource variation generates positive correlations between pre- and postcopulatory sexual traits. Behavioral Ecology, 56, 438–7.

This is a fascinating manuscript, thank you. This is nicely relevant to some of our forthcoming papers.