Dear Recommender,
Thanks for all the comments, from you and the two reviewers. All were very helpful and contributed greatly to improved the paper. Please find below specific answers to each of them. When a page number is indicated, it corresponds to the version without track changes.
Sincerely yours,
Michel Raymond

==

by Jacqui Shykoff, 24 Apr 2022 21:39
Manuscript: **https://doi.org/10.1101/2022.02.22.481477**

request for a revision of the preprint "Increased birth rank of homosexual males: disentangling the older brother effect and sexual antagonism hypothesis"

Dear Authors,

Your very interesting preprint has now been read and assessed by two independent peers who both point out a lot of merit and value in the work. However, both suggest a number of changes that would make this work stronger, clearer and more valuable to the scientific community.

Your reviewers have done an excellent job and provide several useful, constructive suggestions. I have also annotated a copy of your preprint with highlighting and comments that I hope will be easily interpretable. In many places I suggest alternative wording but I also pose a number of questions and point out what I perceive to be inconsistencies. If there is any problem for you in figuring out what I am asking you to do I will provide you with a numbered list, though the easiest for me would be if you could provide me with an editable text in open office or word.

Thanks, all the comments and suggestions have been taken into account. Comments/questions are copied below with their corresponding answer. Other minor suggestions for corrections have all been incorporated and are not reported here.

Page 3, 1st paragraph: “how strong is the fertility cost? Has there been a fertility cost for a long time?”

Quantitative estimates of the cost in various studies are now provided, including a 100% fertility cost reported in a non-industrialized country (Samoa). This suggests that this cost was substantially high for some time, although this temporal aspect is not discussed in this manuscript.

Page 3, 2ndt paragraph, 3rd sentence: “a bit banal as a statement. Of course this is the case.”

This sentence has been removed.

Page 3, end of 2ndt paragraph: “what is this? "fertility on the maternal side"? Is this "higher fecundity of female relatives of homosexual men..."

Yes, this is a higher fecundity of female relatives, but only on the maternal side (e.g. maternal aunts, but not paternal aunts). The sentence has been modified for clarity.

Page 3, 3rd paragraph, 2nd sentence. " 'Family determinants': this could imply genetic or environmental effects. Do you mean "Local environmental effects within the family" or something like that?

Yes, this is now changed

Page 4. line 4. 'replicated': you mean "it has been found repeatedly independently ..."

Yes, this has been changed

Page 4. 1st paragraph: 'typical': Can this be interpreted as a value judgement? You have not written "normal" but ...

The Webster's dictionary definition is: "Typical: exhibiting the essential characteristics of a group". As the groups referred to here are well defined (brain structure of heterosexual males and females), we believe the usage of this word is correct in this sentence, and of course does not imply a value judgment from us. Yes, the word 'normal" would have been inappropriate here.

Page 4. end of 1st paragraph. What two lines of research do you mean? Antagonistic pleiotropy versus OBE proximate mechanisms? It might be worth mentioning that these are two different kinds of mechanisms - evolutionary versus proximate mechanisms. that they are non-mutually exclusive, etc...

Yes, this is now detailed in the new version.

page 5, last paragraph. OK, but what it this biological phenomenon? AP supposes the existance of alleles that favour female fecundity while influencing male sexual orientation. That will, of course, be detectable as an OBE. The OBE will only be visible in large families, i.e., families with mothers with high fecundity but does not rely on the presence of alleles with AP. So not really the same biological phenomenon. i.e., an OBE does not imply AP alleles but AP will generate an OBE and OSE.

Of course, we agree with the referee. This sentence (which was confusing) has been modified, and the ambiguous term 'biological phenomenon' has been removed. Thanks for pointing this out.

Page 6. 1st paragraph. Yes but not the same underlying mechanisms. AP will generate an OBE but OBE does not imply AP.

Yes, the two mechanisms are independent, the confusing wording on "the same biological phenomenon" at the beginning of this paragraph has been removed (see previous point).

Page 6. end of 1st paragraph. *Does this really address the "causality" problem? Maybe it does - I need to think about this.* Beginning of 2nd paragraph. *Does this actually resolve the dilemma? Individuals that produce many children probably DO have alleles for high female fecundity, but those that produce few children may also have these alleles.*

By definition, individuals bearing alleles for high females fecundity have on average a higher fecundity… So, while some individuals with these alleles have a low fecundity, and some individuals without these alleles have a high fecundity, there must be a correlation between having alleles for high female fecundity and having more children than average. Thus, at equal fertility, variation of mean birth rank of homosexual males cannot be attributed to AP.

Page 7, 2nd paragraph. But some of these metrics will cause this, since in sibships of one or two the number of older brothers can only vary between 0 and 1.

These meta-analyses are done on non-restricted data, thus this problem does not arise here. This is now specified in the text

Page 7, 2nd paragraph. Is "index ration" the same as "odds-ratio"?

Odds-ratio are a type of index ratio, but not all index ratio are odd-ratio. “Index ratio” is not consensual, and was removed. See answer to a similar comment by reviewer 2 (this document, page 14).

Page 7, last paragraph. This does not make sense to me. If you use the fecundity of maternal aunts as a measure of female fecundity of the family then you are assuming that there IS a correlation between the fecundity of the mother and her sisters.

Yes, this is true. The sentence has been reformulated accordingly.

Page 8, 2nd paragraph. This statement is a bit overextended. Both OBE and AE are hypotheses proposed to explain patterns of occurrence of male homosexuality within and among families. These are proposed mechanisms that provide predictions that may be consistent with observations. Finding that observations are consistent with the predictions generated by one or both hypotheses would still not constitute evidence that one or both mechanisms were "contributing to the occurrence of male homosexuality". As I understand the point of this study, it is to try to determine whether observations are similarly consistent with predictions from both, indicating that one cannot distinguish one of the hypotheses as more likely, or whether it is possible to distinguish between the two by applying new statistical methods. Whatever you find this will not tell you whether one or both of these proposed mechanisms contributes to the occurrence of male homosexuality ...

The sentence has been rewritten, and there is no more allusion to the idea that one or both mechanisms were "contributing to the occurrence of male homosexuality".

Page 8, 2nd paragraph. I really think that only an experiment (that is impossible to carry out) could answer this question about a causal link between number of older sisters and the probability of being homosexual.

We agree that experiments are impossible to carry. This is why other tools should be developed…

Page 12, last paragraph. Data from 800 individuals classified into three categories were retained: ... (is this correct?)

No, one ambiguous male category was not considered, this is now clearly stated.

Page 12, last paragraph (and also page 13, last paragraph). Were these data not available for the Greek sample? The way you word this maternal aunt fecundity is not available for the Greek sample.

This is correct, maternal aunt data were not available for the Greek data. This is now indicated in the section “Statistical analysis of individual family data”

Page 13, last paragraph. You are using this as the explanatory variable to test whether males of different sexual orientation have more or fewer cousins by maternal aunts.

Yes, this is now clearly stated

Page 13, last paragraph. What, exactly, is a "confounding variable" in comparison with a "control variable"?

No differences. For the model description, ‘confounding variable’ was replaced with ‘control variable’.

page 14, top. These are your (qualitative) explanatory variables? i.e., male sexual orientation? plus the "control" and "confounding" variables? Please be more explicit.

Yes, this now clearly stated.

Page 21, last paragraph:

-Is this the mean (a normal arithmetic mean?) calculated over all data?;

-And this is a subgroup you defined because they had high fertility?

-What does this "mainly homosexuals" mean? You propose AE, i.e. homosexuality arising in families with high female fertility, as a mechanism that can explain persistence of homosexuality in populations despite it being counter-selected. What is the AE for non-homosexuals?

This section has been completely rewritten for better clarity. It should now be more apparent how the values presented are derived from the estimated parameters, and that the subgroups are defined by latent variables during the MCMC chain process. The corresponding table 3 has also been edited for better concordance with the names of the parameters.

page 25. Middle paragraph

number of male offspring per sibship (this is what you mean, right?) or maybe just "such that males have no (or few) brothers"

Actually it CANNOT be very different because there are no brothers. For sibships with only one male the male birth rank must be identical for homo- and heterosexual males.

Yes, this is the intended meaning. We have changed this to make it clearer.

I note that both reviewers suggest an independent assessment of the statistical methods by an expert, and reviewer 2 pleads for more clarity and transparency in your presentation of the modelling and analyses. Because I have already spent far too long before getting back to you with a decision (we just moved our lab and there have been a number of things keeping me busy) I wanted to get this work back into your hands for a thorough revision of the text and presentation.

Thanks !

At the same time, I will look for an appropriate expert in these analytical methods to provide you with more precise suggestions on how to present that part of the manuscript with better clarity.

Already, I think, the comments and suggestions of the two reviewers will guide you in improving all aspects of this work including the model and the analytical methods.

Yes, all the comments are indeed very useful.

I add here a final comment to accompany the notes in the attached file:

The statstical/analytical problem you set out to ellucidate seems to me to be an example of a not uncommon problem of colinearity. A Fraternal Birth Order Effect that will be increasingly visible in large sibships and a maternal fertility effect will, on average, generate a Fraternal Birth Order Effect because you will only find males of high fraternal birth order in large sibships. Now I recognise that introducing additional specific terminology (can we call this jargon) does not usually increase clarity or enhance understanding, but colinearity is a common problem since explanatory variables are often correlated in observational data (size often increases with age, ecological and phylogenetic distance are often correlated, geographic and genetic distances are often correlated ...). As I read your paper I thought "hmmm ... another colinearity problem", so perhaps, and just perhaps, it might be useful to use this framework of trying to distinguish which underlying variable is more important within a context of colinearity. One feature of colinearity is that explanatory power of a variable depends on the order in which it is entered into a sequential model that includes the other correlated variable or, in a simultaneous model, the presence or absence of the other variable. This is, more or less, what you show in Figure 4, so you are already addressing the issue. Perhaps the "colinearity framework" could contribute to the clarity requested by reviewer 2.

Yes, of course, very true indeed. Two points though: 1) Here, while there is indeed a correlation between the fraternal birth order and the fertility (large birth ranks come from larger families on average), importantly, there is A LOT of variation in fertility for low birth ranks: first born, for example, come from families with very variables size. This allows us to test directly for the AP hypothesis independently of the OBE: for males of birth rank 1, the AP hypothesis predicts that homosexuals have more sibs than heterosexuals, while the OBE does not. We thus believe that we have enough power to ensure that a model where birth ranks is controlled for (ie assuming that the OBE exists) allows testing for the AP hypothesis. 2) We are not confident that adding yet another terminological framework would help increase the clarity of the paper.

Yet, as many readers might think the same way as you do, we have added a sentence to the introduction (page 8) to explicitly mention collinearity in the context of our work.

Please revise your manuscript taking into account the comments of all reviewers.

If you have any questions about my comments (I am not very good at annotating pdf's), just ask, and I will try to obtain the opinion of a statistically savy peer rapidly.

sincerely,

Jacqui Shykoff

Reviewer 1

INVITED REVIEW of Raymond et al., "Increased birth rank of homosexual males:Disentangling the older brother effect and sexual antagonism hypothesis." Prepared by Ray Blanchard, Ph.D., Department of Psychiatry, University of Toronto
It is my pleasure to accept the invitation from bioRxiv to review the preprint by Raymond et al. entitled "Increased birth rank of homosexual males: Disentangling the older brother effect and sexual antagonism hypothesis." I am not personally acquainted with Professor Raymond or any of the other authors and I am comfortable that I can review the document impartially. I had a brief email exchange with Professor Raymond some time ago, in which he asked me for a copy of a data set (which I provided), and that is basically all the interaction I remember.

Yes, that is correct.

Studies of samples consisting of homosexual and heterosexual men have found that homosexual subjects tend to have more older brothers and to have larger families (i.e., larger sibships). These associations could arise in different ways: (1) Homosexual men have more older brothers because they come from larger families, or (2) Homosexual men have larger families because they have more older brothers. It is also possible that (3) these two observed associations arise independently from the operation of separate underlying mechanisms. The primary goal of Raymond et al.'s study is to determine which of these possibilities is most likely.

TERMINOLOGY
The association between higher numbers of older brothers and higher odds of homosexuality has sometimes been called the FBOE (fraternal birth order effect); the present authors called it the older brother effect (OBE). The association between larger family sizes and higher odds of homosexuality has sometimes been called the female fecundity effect (FFE); the present authors call it the sex-antagonistic effect (AE). It might be noted that these labels have different relations to the phenomena that they denote. The terms FBOE, OBE, and FFE are purely descriptive and theoretically neutral. In contrast, the term AE is related to the theory that the decreased fertility of homosexual men is offset by an increased fertility in their female relatives (including their mothers). It is possible that this theory arose as an auxiliary hypothesis to reconcile findings of a probable genetic influence on homosexuality with the observation that homosexuality confers a reproductive disadvantage. It would have been my preference that the authors used the relatively atheoretical term FFE rather than AE for the sake of conceptual consistency. I do not, however, feel strongly enough to recommend that that the authors search-and-replace all instances of AE with FFE, because this might entail identifying and re-writing additional relevant sentences in the text, and these could easily be missed in the re-write of a longish manuscript.

OBE has been replaced by FBOE in the edited version (this point was also pointed by reviewer #2). However, as we are indeed addressing the hypothesis that « *the decreased fertility of homosexual men is offset by an increased fertility in their female relatives*", we maintained the use of AE except in some parts where we replaced it with FFE when it was more correct.

SIGNIFICANCE OF THE RESEARCH
This study is important to research on the developmental origins of sexual orientation for at least two related reasons. First, the FBOE/OBE and FFE/AE are potential clues to biological influences on sexual orientation. They could lead – and, in the case of the FBOE/OBE, already have led – to laboratory research on the origins of homosexuality (Bogaert et al., 2018). It is therefore important to establish with survey data which, if either, of these phenomena is reliable and genuine (i.e., non-artifactual), because laboratory research is relatively expensive, labor-intensive, and time-

consuming. One would expect laboratory scientists to be more willing to undertake research on the biological underpinnings of these phenomena if the phenomena themselves have been shown to be reproducible. Second, statistically disentangling parameters of sibship composition (including the OBE and the AE) is difficult and fraught with potential hidden problems. It is therefore desirable to approach such analyses with a variety of different statistical methods, each based on its own sound mathematical-statistical logic, with the goal of obtaining convergent conclusions. The present authors have developed such a set of novel statistical tools to investigate the reliability of the OBE and the AE and to investigate whether one is an artifact of the other.

Thanks for those comments.

The remainder of my review will consist of more specific comments and suggestions. The manuscript appears to have been carefully prepared and well-reasoned. My relevant expertise is in the content area of OBE research rather than mathematical statistics, and this will be reflected in the topics on which I offer comments or suggestions. Some of my comments are in the nature of opinions or asides rather than specific suggestions for additions, deletions, or modifications to the manuscript. I recognize that practically all authors do some amount of cherry-picking the data that they cite or the conclusions that can be drawn from them. This is often necessary for clear exposition and does not necessarily imply any deceitfulness on the part of the author. Thus, I do not expect the authors to pick up on everything that I write, but there are some observations that I would like to make somewhere, and this review presents an opportunity for doing that.

Thanks!

COMMENTS ON THE INTRODUCTION
On page 4, the authors write, "it is still unclear whether the OBE is universal. The OBE is not always found, even in some large samples."
It is easy to make too much of failures to detect an OBE in specific individual samples, even large ones. The OBE is a weak effect (in the statistical sense) perhaps because it shares influence on sexual orientation with numerous other factors, including multiple inherited genes. Thus, researchers will have low statistical power to detect it.

Yes, we agree. In the discussion, we now further discuss this point, particularly for studies from populations displaying a relatively low fecundity.

There is another, different reason that researchers should go into OBE research with the assumption of low statistical power. This relates to the maternal immune hypothesis (MIH), the notion that the OBE reflects the progressive immunization of some mothers to Y-linked antigen by each succeeding male fetus and the concomitantly increasing effects of anti-male antibody on sexual differentiation in the brain in each succeeding male fetus. On this view, live-born older brothers are most likely a proxy for maternal immunization to Y-linked antigens of fetal origin. If the MIH is essentially correct, then older brothers may be quite fallible as a proxy. One does not know how many live-born sons did not expose their mothers to immunogenic amounts of Y-linked antigen, and one does not know how many miscarried (and perhaps never detected) male fetuses did expose their mothers to Y-linked antigen. If any authors want to make the point that the universality of the OBE has not been fully demonstrated, they might do better to point out that nearly all the research has been carried out on subjects of European or Asian descent. There is little research on subjects of sub-Saharan African descent.

Thanks, this interesting point has been added in the discussion (edited version page 27)

On pages 4–5, the authors also write, "the OBE is sometimes described from samples which are not comparable. For example, several meta-analyses . . . attempting to demonstrate an OBE in homosexual men across multiple studies include data from transexuals, pedophiles, hebephiles, or gender-dysphoria individuals . . . . As these different situations are drawn from highly non-representative populations . . . and are not necessarily the result of similar determinants as those for homosexuality, or could represent extreme values from a continuum, considering them could introduce some biases." Writers who make this criticism never point out that the compared homosexual and heterosexual groups were matched on the variable that distinguished them from standard samples; that is, homosexual pedophiles were compared with heterosexual pedophiles, homosexual transsexuals were compared with heterosexual transsexuals, and so on. In general, the homosexual group has been shown to report more older brothers than its heterosexual control group. Thus, one could just as easily cite the findings from special groups as evidence for the broad applicability of the OBE. In practice, the atypicality of the groups is generally given as a reason for dismissing the findings out of hand, either because the writer wants to argue for the need of more research or because the writer would like to dismiss all evidence for the OBE in any way that he or she can.

Thanks for this interesting point. The aim of this paper is first to evaluate the existence of FBOE in homosexual men, controlling for the known biases. Determinants for paedophilia are not identified, thus it is also not known if a homosexual paedophile is just the conjunction of homosexual and paedophile determinants, or if some interaction is at play. The same reasoning applies for the other categories, which are all rather understudied and infrequent. It is thus just safe (and conservative) to restrict the analysis to sample not related to these categories, to avoid criticism on these points. Why the data are restricted as such is now better explained in the text (see page 10). In addition, knowing that the FBOE is validated for typical homosexuals (as this work concludes), it is interesting to confirm that FBOE applies also for these specific categories. This is now mentioned in the discussion (see page 25).

COMMENTS ON THE METHOD AND RESULTS

Raymond et al. have thoroughly explicated the statistics they used (or developed) for their study. As I previously indicated, I will not attempt to critique their methodological innovations, because I have no particular expertise in mathematical statistics and there are plenty of other people who do. I can, however, compare their findings with those obtained in other, recently published analyses of homosexual and heterosexual males. Prior to 2021, a variety of statistical approaches had been used in studies of the FB OE/OBE and FFE/AE. An emerging standard for investigating the FBOE/OBE was a logistic regression model, in which the criterion variable was the subject's sexual orientation, dichotomously coded as heterosexual or homosexual, and the predictor variables were the subject's numbers of older brothers, older sisters, younger brothers, and younger sisters. This approach had some drawbacks and limitations, which caused Blanchard and Lippa (2021) and Ablaza et al. (2022) to develop new statistical procedures, which were different from each other and also different from the procedures used by Raymond et al. Because these studies used completely different methods to approach the same questions, it is useful to consider how Raymond et al.'s conclusions compare with those of the other two studies. Raymond et al. agree with Blanchard and Lippa and Ablaza et al. in two ways: All three studies found positive evidence of an OBE (FBOE) and no evidence for an AE (FFE). Raymond et al. disagree with Blanchard and Lippa and Ablaza et al. in one way: Raymond et al. found no evidence for an OSE (SBOE), whereas both Blanchard and Lippa and Ablaza et al. did find evidence for such an effect. Thus, the study by Raymond et al. helps to clarify which topics within this research area most require further investigation.

Thanks for these comments.

COMMENTS ON THE DISCUSSION
I do not understand some of the remarks by Raymond et al. regarding prior research on the OSE (called the SBOE by other authors), that is, the possible existence of a positive correlation between a subject's number of live-born older sisters and his odds of being homosexual. On pages 23–24, Raymond et al. wrote, "Thus, the report, in a recent meta-analysis, of a widespread OSE in addition to the OBE (Blanchard and Lippa, 2020), should be treated with caution, as we have shown how an apparent OSE is generated when only OBE is acting (fig. 1). . . . This sampling bias does not rule out the action of a genuine OSE in population data, but any claim for an OSE, or for any additional sibling effect, should first control for the sampling bias generated by OBE. . . . We thus conclude that there is, to date, no conclusive support for an OSE in empirical data. The only exception is perhaps from Ablaza et al. (2022), although their new regression method, using several highly correlated variables, requires a formal validation."
In the first place, the study by Blanchard and Lippa (2021) was not a meta-analysis but a reanalysis of a single large data set. It is possible that Raymond et al. are confusing the study by Blanchard and Lippa (2021) with the study by Blanchard, Beier, Gómez Jiménez, Grundmann, Krupp, Semenyna, and Vasey (2021), which did use meta-analysis, and which also found evidence for an OSE/SBOE, albeit using a completely different methodology.

Yes, the correct reference in this sentence is indeed Blanchard et al. 2021 and not Blanchard and Lippa 2021, thanks for pointing out this mistake, this is now corrected.

It is possible that Blanchard and Lippa's finding of an OSE is wrong but it is not possible that it is artifactual. In other words, it is possible that Blanchard and Lippa's finding was simply the result of sampling error (virtually always and everywhere a possibility) but it is not possible that the observed correlation between live-born older sisters and homosexuality was the artifactual result of a correlation between live-born older brothers and homosexuality. That is because Blanchard and Lippa's investigation of the OSE was carried out on a subset of subjects who had no older brothers. Blanchard and Lippa (2021) reasoned that mothers can be immunized by male fetuses, whether these fetuses are subsequently delivered as live-born infant boys or are miscarried. They further reasoned that the number of live-born girls a man's mother delivered before him should correlate with the number of male fetuses that she miscarried before him. Thus, one should observe very slightly higher rates of homosexuality in men who had one or more older sisters but no older brothers compared with men who had no older siblings of either sex. That is exactly what Blanchard and Lippa (2021) found. As far as I can see, the research design of Blanchard and Lippa (2021) is pretty clean, and the results cannot be explained as an artifact of older brothers, because none of the subjects had any older brothers. It is noteworthy that they found that one older sister increased the odds of homosexuality by 12% and two or more older sisters increased the odds of homosexuality by 39%.

It is correct that FBOE is controlled for in some tests presented by Blanchard & Lippa 2021. However, the evidence for SBOE in this work might not be as strong as it is claimed, in our opinion. The most rigorous test is the comparison of groups 2 and 3 (from their Table 1), where family size is the same (thus fecundity is controlled for), and the birth rank differs: 1 (thus no older sib) and 2 (one older sister). In this case, the P-value is 0.02, as reported in the paper, and the conclusion is that having an older sister increases the odd of male homosexuality. However, the younger sib of group 3 is either a male or a female. It would have been preferable to restrict this situation to only a female, thus only the birth order change between the groups, and not also the sex

of the other sib (e.g. female of older sib, and male or female for younger sib). A crude calculation (dividing by 2 the group 3, thus assuming a 1:1 sex-ratio) provides a test for SBOE with P = 0.04. The absence of SBOE is still significantly rejected, but this is not a strong result with such p-value. In the same paper, a replication test is provided with larger families, but fecundity is not fully controlled for, and the group with older sisters (i.e. group 6) displays a larger mean family size than the group without older sisters (group 7), see their table 2: the opposite situation would have been conservative.
In conclusion, we agree that Blanchard and Lippa (2021) present evidence supporting the existence of a SBOE, even if the evidence is not as strong as they suggest, in our opinion. The paper of Blanchard & Lippa (2021) is now discussed in details in the revised version, see page 24, and Supplementary Materials.

Furthermore, the one subsequent published study that also made a preplanned attempt to detect an OSE did detect one (Ablaza et al., 2022) and the magnitude of the OSE was similar. In summary, I think that Raymond et al. should refine their conclusion, in the Abstract, for example, that "An OSE seems to result from a sampling bias in presence of OBE, and is likely to be artefactual." There is little question that the correlation between live-born older brothers and homosexuality does produce an artifactual correlation between live-born older sisters and homosexuality, but Raymond et al. have been too quick to dismiss evidence from other studies that a small correlation between live-born older sisters and homosexuality may remain after one has excluded the effect of live-born older brothers.

We agree and we have modified our text discussing the SBOE to better reflect the findings of Blanchard & Lippa (2021) and Ablaza et al., (2022). The sentence in the introduction concerning SBOE has been changed. The discussion on SBOE (largely edited) ends by: "*We thus agree that there is probably a genuine albeit very small SBOE in empirical data, although a thorough validation is required.* "*(see page 24)*

On page 23, the authors wrote, with regard to prior research, that "the generation of various ad hoc statistics to test various hypotheses has added to the confusion." I think the authors should be careful about throwing around "ad hoc" as a pejorative. One could argue that the statistical procedures they developed for their own study are just as "ad hoc."

All mentions of "ad hoc" have been removed.

On page 27, as elsewhere in the manuscript, the authors state their important conclusion that "After controlling for the confounding effect of the OBE on fertility in families of heterosexuals and homosexuals, we have found no direct association between higher maternal fertility and male homosexual orientation, i.e. no support for genetic factors increasing fertility of females and increasing at the same time the probability that any given son is homosexual." This is an important conclusion, because the balancing selection hypothesis of homosexuality has long been a favorite of evolutionary psychologists, evolutionary biologists, and geneticists. I think it adds to their study rather than detracting from it to point out that other authors have also reached the conclusion that homosexual men do not come from larger sibships than heterosexual men when birth order is taken into account. The study by Blanchard (2012), while not particularly sophisticated with regard to methodology, included the historically important interview data collected by Alfred Kinsey and his associates. The conclusion reached by Blanchard (2012) was not supported by Rieger et al. (2012), but it was supported by Ablaza et al. (2022).

Blanchard (2012) and Rieger et la. (2012) are now cited (see page 8).

MISCELLANY
I have two comments regarding Table 1. First — and this is important — the group labels are reversed for the homosexual and heterosexual subjects in the study by Blanchard and Lippa (2007).

Thanks, this is now corrected.

Second, In Table 1, the data for the homosexual and heterosexual subjects in each study are given in random order: Sometimes the data for the homosexuals are given first and sometimes the data for the heterosexuals are given first. I think it would be highly desirable to make this consistent between studies.

Thanks, this is now corrected.

The section headed "Aggregated family data" on page 19 could be made easier to follow. One has to go back to page 11 to see that male birth rank is computed as OB/N + 1 and female birth rank is computed as OS/N + 1. Reminders of these formulas could easily be given in parentheses on page 19.

The formula is now reported on page 19 for an easier reading of this section.

Similarly, it might be helpful to remind the reader somewhere that the X-axes in Figure 2 are actually Fertility/2.

This is now clearly indicated in the legend of Figure 2.

On page 8, Raymond et al. state "In a meta-analysis, when data are restricted to families with only one or only two sons, no AE is found (Blanchard et al., 2020b), although this analysis includes a paper retracted since then." They need to report what paper was retracted. (I haven't heard anything about it.) This is surely not a secret; there is a whole website called Retraction Watch. This is important information for future authors who need to review the literature or perform meta-analyses on it.

This is the paper of Khorashad et al. However, this retraction was temporary, as its status is now normal. This can be seen on Retraction Watch, see
http://retractiondatabase.org/RetractionSearch.aspx#?auth%3dKhorashad%252c%2bBehzad%2bS
The mention of this retraction has been removed in the revised version, as it is now obsolete.

REFERENCES

Ablaza, C., Kabátek, J., & Perales, F. (2022). Are sibship characteristics predictive of same sex marriage? An examination of fraternal birth order and female fecundity effects in population-level administrative data from the Netherlands. Journal of Sex Research. https://doi.org/10.1080/00224499.2021.1974330

Blanchard, R. (2012). Fertility in the mothers of firstborn homosexual and heterosexual men. Archives of Sexual Behavior, 41, 551–556. https://doi.org/10.1007/s10508-011-9888-0

Blanchard, R., Beier, K. M., Gómez Jiménez, F. R., Grundmann, D., Krupp, J., Semenyna, S.W., & Vasey, P. L. (2021). Meta-analyses of fraternal and sororal birth order effects in homosexual

pedophiles, hebephiles, and teleiophiles. Archives of Sexual Behavior, 50, 779–796. https://doi.org/10.1007/s10508-020-01819-3
Blanchard, R., & Lippa, R. A. (2007). Birth order, sibling sex ratio, handedness, and sexual orientation of male and female participants in a BBC Internet research project. Archives of Sexual Behavior, 36 163–176. https://doi.org/10.1007/s10508-006-9159-7
Blanchard, R., & Lippa, R. A. (2021). Reassessing the effect of older sisters on sexual orientation in men. Archives of Sexual Behavior, 50, 797–805. https://doi.org/10.1007/s10508-020-01840-6
*Bogaert, A. F., Skorska, M. N., Wang, C., Gabrie, J., MacNeil, A. J., Hoffarth, M. R.,VanderLaan, D. P., Zucker, K. J., & Blanchard, R. (2018). Male homosexuality and maternal immune responsivity to the Y-linked protein NLGN4Y. Proceedings of the National Academy of Sciences of the United States of America, 115, 302–306.*https://doi.org/10.1073/pnas.1705895114
Raymond, M., Turek, D., Durand, V., Nila, S., Suryobroto, B., Vadez, J., Barthes, J., Apostoulou, M., & Crochet, P-A. (2022). Increased birth rank of homosexual males:Disentangling the older brother effect and sexual antagonism hypothesis. bioRxiv. https://doi.org/10.1101/2022.02.22.481477
Rieger, G., Blanchard, R., Schwartz, G., Bailey, J. M., & Sanders, A. R. (2012). Further data concerning Blanchard's (2011) "Fertility in the mothers of firstborn homosexual and heterosexual men" [Letter to the Editor]. Archives of Sexual Behavior, 41, 529–531. https://doi.org/10.1007/s10508-012-9942-6

Reviewer 2

Comments for Authors

In the present manuscript, the authors seek to examine two important questions for understanding the evolution of male same-sex attraction. First, some work has shown that the female relatives of homosexually attracted males have higher reproductive output than the female relatives of heterosexually attracted males. This has been interpreted as suggesting one of several potential genetic effects that could account for the existence of male homosexuality, given that genes associated with this attraction pattern in males might lead to elevated reproduction when carried by females. Alternatively, the fraternal birth order effect (FBOE), or as the authors prefer the Older Brother Effect (OBE), is the finding that each older brother a male has increases the likelihood that he will have exclusive same-sex attractions. Few attempts have been made to disentangle whether the elevated reproduction of female relatives due to some separate genetic effect or simply a result of highly fecund women being more likely to produce homosexual sons. The work is timely, interesting, and a solution to this empirical problem would be welcomed by many. Although I find much to praise in the manuscript, and think the authors' desire to bring innovative solutions to this challenging problem is laudable, I cannot recommend the manuscript in its present form. My largest concern, as detailed below, revolves around ease of understanding the proposed analytic models. Many analytic models are difficult to understand and this should not be avoided if they nonetheless improve accuracy and understanding of phenomena. However, all analytic frameworks must eventually be implemented, and I think more researchers in this area would be likely to implement the proposed frameworks if they were easier to understand. More specific comments are found below, and I hope they are of use to the authors. I similarly hope that my focus on areas of concern does not overshadow my largely positive impression of the manuscript—these comments are offered in the hope that the manuscript can be refined and hence made more useful to the scientific community who will pay attention to such scholarship.

Thanks for these positive comments. More details have been added to facilitate the use of the new tools presented here for other scientists.

1) Abstract: I'm not sure I know what the authors mean by "evolutionary determinants." This leaves quite a bit of ambiguity as to whether the considerations are proximate or ultimate in nature.

Evolutionary determinants are definitively 'ultimate'. This is now specified in the revised version (see abstract).

2) Abstract: It's unclear why an antagonistic effect (AE) would lead to a discernible birth rank effect rather than simply a greater number of siblings in general.

The sentence has been changed, and the link between AE and higher mean birth rank is now clearly stated. This link is also explained in more details throughout the MS.

3) The first sentence of the introduction is rather lengthy, and can be a difficult compound sentence to parse. I might recommend splitting this into two sentences to reduce reader burden.

Thanks, this has been done.

4) When discussing the origin of male homosexuality, it would be helpful if the readers distinguished between proximate origins (e.g., certain genes, maternal immune factors, etc.) as opposed to ultimate explanations for the origins of the trait (balancing selection, spandrel, etc.). These explanations are of course enmeshed, but clarity would be appreciated.

In the introduction, the text has been clarified to specify when a proximate (FBOE) or an ultimate (AE) mechanisms is discussed. This is further discussed in the Discussion (see sections *Proximate and ultimate mechanisms of the older brother effect*, and *No support for antagonistic pleiotropy through female fertility*).

5) The Fraternal Birth Order Effect (FBOE) is typically referred to with the accompanying acronym. Would the authors be open to modifying the recurring acronym OBE (older brother effect) for the FBOE to maintain slightly more consistency in this literature?

Yes, this has been changed throughout.

6) Page 4 & 5. The authors point out that the FBOE has been examined in a huge diversity of samples. This point is well taken, but I would hasten to point out that each of these samples share key developmental features and findings—the FBOE is found among males who are androphilic, and this pattern emerges irrespective of other factors such as gender identity or age-preferences. Although other developmental factors likely play into the variations discussed, these groups all share a remarkably similar FBOE as it pertains to the development of male androphilia. A more relevant caveat in this literature are stopping rules (e.g., Blanchard 2022; https://doi.org/10.1080/00224499.2021.1984379)

Thanks for pointing out this point. The caveat of stopping rules is now mentioned in the revised version. (Page 30).

7) Page 5, discussion of the OSE. Just a note to say that I thought this paragraph was well articulated. Sometimes as reviewers we get mired in finding areas for potential improvement, but some praise also seems warranted.

Thanks for this comment!

8) Page 6. The authors state: "Indeed, a higher fertility of mother of homosexual men implies that, when sampling homosexuals from a population, the mean birth order of homosexuals is higher than the mean birth order of heterosexuals." I'm not sure if I agree with the logic here. The high fertility doesn't speak to birth order per se, but simply family size. The AE would imply that family size of homosexual males will be large irrespective of actual birth order. The OBE, as stated in the next sentence, would imply that mothers with especially large reproductive output would have more (later born) homosexual sons.

When fecundity increases, the mean birth rank increases in a random sample (random relatively to birth rank), as explained in the Result section "Relationships between mean birth rank and mean fertility". The mean birth order of homosexual will be higher **on average** (this is now specified in the revised version, see Page 6, top)

9) Page 6. From my understanding of Khovanova (2019, as published in Archives of Sexual Behavior) is that one proposed approach (Method 1) examines restrictive cases of families with either 1 or 2 children. A more relaxed formulation (Method 2), also detailed in that paper, is to ignore sisters and consider families with either 1 or 2 males (which from a data analytic perspective amounts to participants with either no brothers, or one brother). The authors are correct, however, to note that these restrictions can substantially reduce sample sizes.

Thanks for this comment. The paper of Khovanova is now fully published in Archives of Sexual Behavior, in 2020, and the reference has been updated.

10) Page 7. I'm not sure what the authors mean when they say "index ratios." Price and Hare (1969) as cited certainly use the term "ratio," but "index ratio" does not appear in that manuscript.

Index ratio was unfortunate. This sentence has now been amended.

The authors' point about limitations are well taken, but there is a certain "hollowness" in this criticism given that one couldn't reasonably expect to have "matched controls" in the same way as many medical studies and such that employ Odds ratios. Groups of homosexual individuals are already (statistically) unusual, and so removing potential "confounds" that differentiate them from heterosexual individuals may also remove important causative factors.

Yes, this is true, but not considering "confounds" could lead to biased results. A classical example is age: is one group is younger than the other, it might display fewer younger sibs for that reason, and thus eventual results on that sib category could be artefactual. Or if the mean age is the same, but the mean year of birth is different (e.g. sampling was done at different time for both groups), and if mean fertility varies with time (as it does currently in many countries), then the mean birth rank between the samples will vary at least for that reason. And so on. We believe that it remains pivotal to be able to statistically control for those potentially confounding variables.

The authors continue by saying, "Also, using an index ratio implicitly assumes that the expression of OBE is independent of the level of fertility, which remains to be shown." I'm not sure I understand the meaning here. If the MIH is correct, and the FBOE is genuine, then higher fertility will invariably lead to more (later born) sons who are same-sex attracted. This fact, and this fact *alone, may explain related findings that the mothers of homosexual males sometimes show elevated reproduction—they end up in these "samples" precisely because they had high fertility to begin with.*

This sentence has been suppressed.

11) I hope the authors don't take my concerns detailed above as an indication that I do not see merit in their arguments. They have conveyed a key concern in this literature about both theoretical considerations, and the data typically collected, and argued that we need more refined methodological tools to disentangle AE and OBE.

Thanks for this comment !

12) Page 9. I don't think I follow (or agree with) the logic outlined for the twin comparison. The authors state: "under the AE hypothesis, there should thus be more homosexual men among the dizygotic twins than among the monozygotic twins." There is little (if anything) in the introduction that would assist readers in seeing the rationale for this prediction without devoting their own further thinking to the matter. The authors could help lead readers down this path more clearly. I understand that the authors are making a simple "prevalence rate" prediction (i.e., not a concordance rate between twins), but I'm not sure why the presence of an AE would imply this prediction. Tong and Short (1998) discuss the ebb and flow of dizygotic twinning rates across time, which seem to be marked by initial decreases in developed countries, and have increased since the 1980s because of exogenous fertility treatments. Indeed, the abstract of Tong and Short (1998) states: "The rise in the dizygotic twinning rates which occurred from the 1980s onwards in developed countries is almost certainly due to increasing use of ovulation-inducing agents, but this rise may have masked a continuing decline in dizygotic twinning." The authors have outlined an evolutionary rationale, and numerous potential confounds in the extant literature. It is unclear to me why examining the prevalence of homosexuality among dizygotic and monozygotic twins represents a step forward when the elevation in dizygotic twinning observed in recent decades is attributable to exogenous fertility treatments that were unavailable to ancestral women. Indeed, women seeking fertility treatments are arguably more likely to be compensating for suppressed natural fertility. This seems to answer a confound with an even larger confound. Furthermore, greater fertility must be distinguished from actual reproductive output—dizygotic twins may be an indicator of "fertility" as defined by ease of conception, but this does not mean that these women invariably have higher overall reproductive output.

The rational was that dizygotic (and not monozygotic) twins were considered as an expression of a higher fertility (found in large families). And fertile families (including those displaying dizygotic twins) should display more homosexual men (under the hypothesis of AE) than a control (families displaying a monozygotic twins). However, a thorough review on the link between dizygotic twinning and fertility has recently shown that it is a sampling artefact (Rickard et al. 2022). There is thus no reason to continue considering twins, as the initial rational is not valid any more. We have entirely removed the twins analysis from this new version.

Rickard,I, Vullioud, C., Rousset, F., Postma, E., Helle, S. Lummaa, V., Kylli, R., Pettay, J., Roskaft, E., Skjarvo, G., Störmer, C., Voland, E., Waldvogel, D., Courtiol, A. 2022. Mothers with higher twinning propensity had lower fertility in pre-industrial Europe. *Nature Communications* **13,** 2886 (2022). https://doi.org/10.1038/s41467-022-30366-9

13) The authors note many exclusions regarding data in published literature, and I think that many of them are justifiable (or at least understandable). However, the authors explain in the introduction that individuals with gender-dysphoria would be excluded. Given this exclusion, it is curious to me that samples from both Samoa and Southern Mexico (i.e., Paul Vasey's research group) are included. These samples are not comprised of heterosexual and homosexual "men," but instead heterosexual men and third-gender individuals known as either fa'afafine (Samoa) or muxes (Mexico). Although not precisely fitting the characterization of gender dysphoria, these third-gender individuals are not necessarily the same as Western gay men (although they do share in common the fact that they are male, and attracted to men/masculinity, among many other correlates such as the FBOE).

The "third gender" qualification (with specific names such as fa'afafine , waria, hijra, berdache,…) is indeed a cultural attribution in some specific societies. Other societies have other types of cultural qualification (such as 'gay' in occidental countries). However, variable cultural qualifications do not reflect the biological unity displayed by those male individuals, as mentioned by the reviewer, such as same-sex attraction and FBOE. It seems that considering third gender individuals as homosexuals is a common practice. Paul Vasez's research group uses the term "andophilic males" allowing them to study indifferently --and to compare-- fa'afafine, muxe (third gender in Samoa or Mexico, respectively), 'gay' or 'homosexual', (see e.g. VanderLaan & Vasez, 2013, VanderLaan, Petterson, Vasey, 2015, Gomez et al. 2018, etc.). In addition, gender dysphoria (usually defined as the distress a person feels due to a mismatch between their gender identity and their sex assigned at birth) does not applies in particular to third gender individuals (considering the current scientific literature). A definition of gender dysphoria has been added (see page 10) in order to clarify the text concerning the exclusion criteria.

Additionally, I am not sure how convinced I am that including bisexual individuals in the French, Greek, and Indonesian samples is appropriate. Male bisexuality in the West is relatively rare, and the determinants are not well understood (e.g., Bailey et al., 2016; Jabbour et al., 2020; www.pnas.org/cgi/doi/10.1073/pnas.2003631117). It is not theoretically parsimonious to lump bisexual and homosexual men together, and this is not typically done in this literature (and when it does occur, it likely should not). Furthermore, the sample described in Nila et al. (2019), which must actually be traced back to Nila et al. (2018; https://link.springer.com/article/10.1007/s10508-018-1202-y) for full sample information, also includes a substantial number (n = 34) of bisexual men. This aspect of the Nila data is not discussed presently, and is difficult to ascertain for readers unfamiliar with this literature. Beyond this concern, the "homosexual men" group in Nila et al. (2018) contains data from 11 gender-variant waria (from my understanding, these individuals are broadly comparable to the Samoan fa'afafine or the Mexican muxes). This is an even more confusing inclusion give that the authors are at great pains to exclude samples documenting the FBOE among androphilic males who are not cisgender. In short, readers shouldn't have to dig this much for sample information, and the samples considered should also remove the rather easy critique/confound of lumping together males who report bisexuality with those who report exclusive homosexuality.

Lumping bisexuals and homosexuals is indeed a common practice (see e.g. a recent example in Blanchard and Lippa (2021) where they say: *"In the present study, as in the birth order analysis in the earlier one, bisexual probands were combined with homosexual probands").* Is it a good or a bad practice? No specific work has been addressing directly this question in the literature, even if partial and scattered results –in one direction or the other – can be found. Nevertheless, if bisexuals are closer to heterosexuals than to homosexuals for the traits under study, pooling together bisexuals and homosexuals would add some extra "noise" and lower the significance of the test, thus being a

conservative approach but also increasing the chance of finding non-significant results. Thus, in case of a non-significant result, one should consider not pooling bisexuals with homosexuals. In the present study, non-significant results were found for FBOE (French dataset), and for FFE (all individual datasets), see Table 4. Removing the bisexual individuals, for each dataset, did not changed qualitatively these results. This is now indicated in the test, in the “Limits and future directions” section (see page 30). See also page 12.
For the Indonesian dataset described in details by Nila et al. 2018, some men self-declaring bisexuals were in fact hesitating, and after further questioning few of them were either heterosexuals (sex worker men, who regularly have sexual activity with men for financial reasons, as they were unsure if the orientation questions were concerned with their business or personal interest) or rather homosexuals (e.g. married man due to social homophobia, or insecure about disclosing their homosexuality). This process of better understanding this bisexual category is detailed in Nila et al. 2018: remaining bisexuals (not firmly identified as heterosexuals or homosexuals) were studied separately, with the conclusion that they did not differed qualitatively from homosexuals for the traits studied (e.g. age, education, income, number of sibs, number of nieces and nephews, direct an indirect reproduction...). Thus, they were pooled with the homosexual category. The reference to Nila et al. 2018, which contains all the details of the sample, has been added where the Indonesian sample is presented (see page 12). Yes, the waria individuals in the Indonesian sample are representative of the third gender. As stated in the answer to the comment just above, third gender men are “androphilic males”, a synonymous of homosexual men. The fact that the Indonesian sample of homosexuals contains also bisexuals and third gender (waria) is now clearly specified in the text (see page 13).

14) Page 11. The authors compute mean birth order as OB/N + 1. This is not a metric that I have seen previously in this literature. I presume that the capital N indicates that this metric is for the entire sample, although it is arguably best applied within sexual orientation category to compensate for sample size differences across groups that are common in this literature. I have perhaps misunderstood the metric, but if I have this only serves to illustrate that the authors could briefly contextualize on the use of this parameterization and how readers should think about it.

A sampled man displaying n older brothers has a birth rank, with respect to only his brothers, of n+1. The aggregated version is similar: in a sample of men, the mean birth rank of these men, with respect to only their brothers, is $OB/N + 1$, *where* OB/N *is the* mean number of older brothers (where OB is the total number of older brothers, and N is the number of sampled individuals). The explanation of this metric has been clarified in the edited version (see page 11).

15) Page 12. Including a standard deviation to accompany age would be helpful. The authors have presumably used a Wilcoxon test for this group difference due to the skewed nature of the data, but a brief note to this effect would ensure readers do not have to engage in any guess-work.

The standard error of the mean has been added for age (see page 12). The standard deviation is indicated in the table of descriptive statistics (Table 2). Details of the Wilcoxon test are provided.

16) Regarding the statistical models, the Editor (or for this preprint service, the “recommender”) will likely need to have this evaluated by a statistician who can appropriately adjudicate the models. That said, the authors have done little to explicate their approach and explain why it represents a useful innovation over previous approaches. Doing so is vitally important, as most readers curious about the FBOE will not have the requisite mathematical sophistication to evaluate the meaning or utility of the models presented. If the authors believe their approaches are a step forward, it is imperative that the rationale be conveyed to researchers in a clear and *understandable manner. (Ablaza et al., 2022 have done exactly this, giving researchers a clear road-map to mimic their analytic framework for both data analysis and interpretation.) To be frank, after several readings of the mathematical portion of the manuscript, I am not sure what the models represent, how to interpret the values, or how this helps push our understanding forward. This is not helped by the fact that model specifics are found in the (equally confusing) supplementary material, yet the main text makes constant reference to specific model numbers in that supplementary material. This creates an incredibly cumbersome reading process that compounds comprehension difficulties. To put my concerns into concrete terms, when I read Ablaza et al. (2022), I am able to see the way they have set up their models, and which exact parameters I should use if I wanted to replicate their models in my own (or publicly available) data. When I examine Table S1, I am completely at a loss for how I would translate this information into parameterization that I could apply to my own data. The authors may attribute this lack of ability to this reviewer's ignorance, and they may be correct to do so, but I will once again say that most researchers in this area do not have the kind of statistical sophistication to decipher and apply these models in their own work. If the authors believe these analytic approaches are a step forward and would like to see them widely implemented on both past and future data, then they will need to translate practical steps for doing so to us mere mortals.*

Classical methodologies, such as generalized linear regression, are clearly explained (and similarly, it is easy for Ablaza et al. to fully detail their statistical approach, as it requires only a classical regression). Some other methods presented here are indeed complex (such as Bayesian analysis using a hierarchical model), and are not accessible without some specific expertise (mastered by the co-author D. T.). The principle of the analyses has been stated in the text, but all the details of these complex tools cannot be given in the text: instead, relevant references are given, as this is usually done. In addition, all the scripts and data used in this paper are easily accessible (deposited in a public repository, as requested by PCI). Thus, anyone can get all the tools to reproduce the results of this study, and eventually to use them for their own data (although a minimum understanding of Bayesian statistics is required).

17) Regarding some parameters in the model, the authors use a 1:1 Male:Female ratio, whereas this number typically shows a slight bias to males (i.e., ~105 Males for each 100 females born; https://ourworldindata.org/gender-ratio). I understand that the slight adjustments are unlikely to impact the models much, but it is worth either adjusting things slightly, or pointing out these caveats.

This point is now clearly mentioned as a limitation, see page 30.

18) The authors employ WAIC, but do not explain what this statistical approach is. Much like the mathematical proofs and models put forward, the authors seem to have assumed a level of mathematical sophistication that will in fact be quite rare among the readers of this kind of manuscript. Similarly, the authors employ RJMCMC without any explanation of the approach, again presuming a huge amount of reader background knowledge and mathematical sophistication.

These rather technical concepts are presented with relevant references to help reader to go further if they wish to learn more. An explanation of the acronymes WAIC and RjMCMC has been introduced, and a short explanation for each has been added (WAIC: page 14, RJMCMC: page 15).

19) Page 24. The authors state, "First, we filtered out dubious samples, such as samples not corresponding to adult homosexuality, such as pedophiles, or corresponding to non-representative populations, such as sex offenders, transexuals, psychoanalytic or hospital patients (Zietsch, 2018).

Thus, our results can be safely associated with standard homosexual men." In defense of Ray Blanchard's (many) contributions to this literature, I hasten to point out once again that these individuals all share in common the relevant factors for the FBOE—older brothers increasing the likelihood of androphilia among males. These other aspects of sexual orientation or presentation notwithstanding (e.g., age preferences, commission of sexual crimes, etc.), that simple fact remains. Indeed, it is interesting that in such diverse samples, the FBOE for male sexual orientation can be found. Indeed, rather than being "dubious" samples, I think they represent a fascinating convergence of evidence. Last, Paul Vasey's group have argued that other expressions of male androphilia have been common throughout most cultures and history (e.g., VanderLaan et al., 2013; DOI 10.1007/s12110-013-9182-z), meaning that "standard homosexual men" may not be an entirely intelligible way to characterize our understanding of the evolution of male androphilia.

Thanks for this interesting point (also raised by the first reviewer). The aim of this paper is first to evaluate the existence of FBOE in homosexual men, controlling for the known biases. Determinants for paedophilia are not identified, thus it is also not known if an homosexual paedophile is just the conjunction of homosexual and paedophile determinants, or if some interaction is at play. The same reasoning applies for the other categories, which are all rather understudied and infrequent. It is thus just safe (and conservative) to restrict the analysis to sample not related to these categories, to avoid criticism on these points. However, knowing that the FBOE is validated for typical homosexuals (as this work concludes), it is interesting to confirm that FBOE applies also for these specific categories. This is now indicated in the discussion, see page 25. (and the word "dubious" has been removed).

20) Page 24. The authors note that low fecundity could be at play in studies that do not find an OBE, but do not similarly acknowledge (here or elsewhere) the way that stopping rules influence these patterns.

The specific case of stopping rules is now mentioned in the section "Limits and future directions", see page 30.

21) Page 26. The authors state: "Based on current knowledge, this OBE is only found in humans and not in any other species of mammals, even those closely related. This suggests that the OBE is not a mere constraint of the gestation in primates, and thus the interference of male birth order with sexual orientation requires an evolutionary explanation." I'm not sure this point is entirely fair given that 1) exclusive same-sex preference is not found among other primates (although bisexual behavior can be widespread), and 2) to my knowledge, no studies of birth order and sexuality have been undertaken among non-human primates. Domesticated rams are the only species to show exclusive same-sex preference among males. Although a birth order study of these rams would be hugely informative, to my knowledge no such study has ever been conducted.

Yes, it is true that it is difficult to speculate on the phylogenetic distribution of FBOE in primates when same-sex preference is not found among non-human primates. Thus, instead of FBOE, this paragraph discuss now about the more general notion of "effect of birth order on reproduction" (see page 27).

The authors go on to speculate that the OBE could be "an adaptive plastic manipulation of the phenotype of male offspring by the mother." This adaptationist logic has a straightforward developmental rebuttal—the OBE is simply a developmental byproduct (i.e., spandrel; Gould & Lewontin, 1979) of actual adaptive processes (i.e., maternal immune system functioning then colliding with the typical neurodevelopment of a male fetus).

The main candidate for a proximate mechanism for the FBOE is a maternal immune response to male-specific antigens. But this is just a proximate mechanism, and it cannot be used as an ultimate explanation, particularly considering the relative high frequency of homosexual men (associated with a relatively high reproductive cost). The ultimate explanation is not currently known (at least there is actually no consensus), and development constraint are also perhaps to be considered: this paragraph discusses this issue and points for direction for future research. This paragraph has been edited for better clarity, and the last sentence of this paragraph is an abstract of the situation: *"… the relative contribution of adaptive responses and developmental constraints in shaping the FBOE, and the selective pressures generating the FBOE, remain in urgent need of investigation". (see page 27)*

22) The authors' conclusion that the OBE is operating, whereas evidence for the AE and OSE is more limited, are well taken. This conclusion represents a step forward in this literature. However, I would strongly encourage the authors to convey the path taken to this conclusion in a more parsimonious and easy to follow manner.

Thanks for the comments on the general conclusion of this study. All the comments from both reviewers, and those of the editor, have helped to tremendously improve the manuscript at various places. The edited version should be hopefully clearer.