THE ROYAL SOCIETY PUBLISHING

PROCEEDINGS B

Visually attending to a video together facilitates great ape social closeness

Wouter Wolf and Michael Tomasello

Article citation details
Proc. R. Soc. B **286**: 20190488.
http://dx.doi.org/10.1098/rspb.2019.0488

Review timeline

Original submission:	27 February 2019
1st revised submission:	23 April 2019
2nd revised submission:	20 June 2019
Final acceptance:	27 June 2019

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

Review History

RSPB-2019-0488.R0 (Original submission)

Review form: Reviewer 1

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

Comments to the Author
The authors examined if apes affiliate more with a person or conspecific who had watched a video together. I really like their idea, and the MS is written well.
My major concern is the lack of baseline in each measure (latency, proximity). Thus, the causality of effects is unclear. That is, unlike what they discussed, if the subject was already a friend of the experimenter/conspecific, she might have approached faster and spent time together, as well as, watched a video together with the experimenter/conspecific. I recommend the authors to include baseline measures to exclude this possibility. I appreciated that they tried to control a friendship among the two subjects in exp 2, but more stringent or quantitative control (i.e. baseline measures) may be necessary for this purpose.
Minor,

The title "visually attending to things together" implies that the content of the things does not matter. But there is no data supporting this. Videos instead of things?
Page 4. Please describe what was "uninterpretable behaviour".
Page 4. Please describe what was "physiology interfering with the eye tracking set up"
Page 6. Please describe the rationale for using a video of a playing baby chimpanzees, and describe the content in more details (a figure shows a baby was playing with two adults in the wild?). I don't think the following is what you intended, but this reminded me of that in SNS, humans enjoy watching "cute" videos together. But apes are unlikely to do the same. From the observational studies, it probably makes more sense to show outgroup male chimpanzees to male chimpanzees to make their ingroup bonds stronger.
Page 8. I recall that I was recommended by someone to use ICC rather than pearson's r to measure inter-coder reliability because pearson r can be applied to linear data. Anyway, it is not clear if you use ICC or pearson's R.
Page 9. Please use examples for "a particular way that apes in zoos relates to humans based on their extensive experience with them"
Page 12. Same as above. Please describe the rationale for using non-chimpanzee animal video as a stimulus. And please describe in more details the content. The video content may matter for the results.

Review form: Reviewer 2

Recommendation
Accept with minor revision (please list in comments)

Comments to the Author
See attached review (Appendix A).

Decision letter (RSPB-2019-0488.R0)

05-Apr-2019

Dear Dr Wolf:

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

Associate Editor
Board Member: 1
Comments to Author:
This is a neat test of a novel idea, with interesting consistency across contexts (chimpanzee-human, bonobo-human, chimpanzee-chimpanzee). More detail on the methods is required in order to fully evaluate the study. For example on the choice of video stimulus, definition of closeness in video analyses and reasons for the absence of control conditions (baseline observations or both looking at "no video"). These comments from the referees should be addressed in a revision.

As a non-great ape specialist, I fully agree with referee 2's comments on framing the study with the broader, proceedings B audience in mind. I was surprised that bonobo and chimpanzee data were not analysed separately in experiment 1 – a point also raised by referee 2. As a non-specialist, it's difficult to tell whether this, or smaller decisions (e.g. observing for < 3 minutes or actively attracting animals in the approach test), are standard or should be justified. I would recommend a brief comment where appropriate on the training required to use juice and the methods used to move individuals into the experimental arena. A statement on statistical methods for each experiment would help keep track of sample sizes.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
The authors examined if apes affiliate more with a person or conspecific who had watched a video together. I really like their idea, and the MS is written well.
My major concern is the lack of baseline in each measure (latency, proximity). Thus, the causality of effects is unclear. That is, unlike what they discussed, if the subject was already a friend of the experimenter/conspecific, she might have approached faster and spent time together, as well as,

watched a video together with the experimenter/conspecific. I recommend the authors to include baseline measures to exclude this possibility. I appreciated that they tried to control a friendship among the two subjects in exp 2, but more stringent or quantitative control (i.e. baseline measures) may be necessary for this purpose.
Minor,
The title "visually attending to things together" implies that the content of the things does not matter. But there is no data supporting this. Videos instead of things?
Page 4. Please describe what was "uninterpretable behaviour".
Page 4. Please describe what was "physiology interfering with the eye tracking set up"
Page 6. Please describe the rationale for using a video of a playing baby chimpanzees, and describe the content in more details (a figure shows a baby was playing with two adults in the wild?). I don't think the following is what you intended, but this reminded me of that in SNS, humans enjoy watching "cute" videos together. But apes are unlikely to do the same. From the observational studies, it probably makes more sense to show outgroup male chimpanzees to male chimpanzees to make their ingroup bonds stronger.
Page 8. I recall that I was recommended by someone to use ICC rather than pearson's r to measure inter-coder reliability because pearson r can be applied to linear data. Anyway, it is not clear if you use ICC or pearson's R.
Page 9. Please use examples for "a particular way that apes in zoos relates to humans based on their extensive experience with them"
Page 12. Same as above. Please describe the rationale for using non-chimpanzee animal video as a stimulus. And please describe in more details the content. The video content may matter for the results.

Referee: 2

Comments to the Author(s)
See attached review

Author's Response to Decision Letter for (RSPB-2019-0488.R0)

See Appendix B.

RSPB-2019-0488.R1 (Revision)

Review form: Reviewer 1

Recommendation
Accept with minor revision (please list in comments)

Comments to the Author
I appreciate that the authors explained the pros and cons of not including no video condition or baseline. My last comment is to intricate the discussion regarding this issue in the text. For example, the test situation in Study 1 is somewhat reminiscent of the established test in which a

chimp has to request for foods from someone facing toward the chimp (attentive) but not from someone facing elsewhere (not attentive). But Study 2 can kill this alternative because the stooge is a chimp. In Discussion of Study 1, after "that is, ", I don't really understand what they are really trying to say. Consequently, I feel that the rationale for Study 2 is still weak. For the potential species difference, they can at least compare the latency data statistically and show no significant difference between the species. If they found a species difference, it is better to treat the data separately or include the species factor in LMM and present this additional analysis in SI.
In Study 2, as Reviewer 2 said, there may be a possibility for the eye-contact effect. It would be nice if the authors briefly discuss or at least argue against this possibility in General Discussion or somewhere because many readers may wonder the same or similar thing.

Review form: Reviewer 2

Recommendation
Accept as is

Comments to the Author
The authors did a nice job with the revision. Where I still might quibble about the significance of the findings that authors did a nice job discussing in discussion.

Decision letter (RSPB-2019-0488.R1)

17-Jun-2019

Dear Dr Wolf:

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

Associate Editor
Board Member: 1
Comments to Author:
Thank you for your response. Referee 2 is overall satisfied with the revisions, though is not fully persuaded on the significance of the study. As referee 1 indicates, this may in part be because the general discussion doesn't offer much in the way of synthesis of the two studies. Nor does it fully place the results in context (just 3 citations), or discuss limitations of the study that you describe in the response. For example, those relating to the control conditions, where relevant citations of studies making similar or different experimental choices are also required. Please revise this section based on the additional comments of referee 1.

My main concern remains, which is about the inclusion of bonobos. Sample size is not a good justification for lumping together two species. A rationale should include citations to support an evolutionary/behavioural predication that the two species should respond similarly, a methodological note on the latencies you observed, and citations to show precedent for combining these species in other studies that are not specifically setting out to make a cross-species comparison. You've indicated that the sample size is unfortunately too small for a parallel bonobo analysis. I therefore agree with referee 1 that you should either leave out the bonobos or include species in the model and discuss these results. Indeed, this could be very interesting if you find generality not only across the contexts that you've explored but across species!

Finally, thank you for taking care to respond to the referees questions in your response. A note and citations are also required in the text for the child and ape studies upon which you base aspects of your experimental methods and for the sentence "animal behavior used in a previous study with human children". Again, the rationale for 3 minutes instead of 1.5 is weak, and could be considered fishing – some justification from previous papers and/or your own observations would be helpful. "Extensive experience" and "sometimes they were motivated by placing some food in the area" need to be more specific, perhaps with a note in the table per individual/trial.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
I appreciate that the authors explained the pros and cons of not including no video condition or baseline. My last comment is to intricate the discussion regarding this issue in the text. For

example, the test situation in Study 1 is somewhat reminiscent of the established test in which a chimp has to request for foods from someone facing toward the chimp (attentive) but not from someone facing elsewhere (not attentive). But Study 2 can kill this alternative because the stooge is a chimp. In Discussion of Study 1, after "that is, ", I don't really understand what they are really trying to say. Consequently, I feel that the rationale for Study 2 is still weak. For the potential species difference, they can at least compare the latency data statistically and show no significant difference between the species. If they found a species difference, it is better to treat the data separately or include the species factor in LMM and present this additional analysis in SI.
In Study 2, as Reviewer 2 said, there may be a possibility for the eye-contact effect. It would be nice if the authors briefly discuss or at least argue against this possibility in General Discussion or somewhere because many readers may wonder the same or similar thing.

Referee: 2

Comments to the Author(s)
The authors did a nice job with the revision. Where I still might quibble about the significance of the findings that authors did a nice job discussing in discussion.

Author's Response to Decision Letter for (RSPB-2019-0488.R1)

See Appendix C.

Decision letter (RSPB-2019-0488.R2)

27-Jun-2019

Dear Dr Wolf

I am pleased to inform you that your manuscript entitled "Visually attending to a video together facilitates great ape social closeness" has been accepted for publication in Proceedings B. Congratulations!!

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

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

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

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

Your article has been estimated as being 9 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Paper charges
An e-mail request for payment of any related charges will be sent out after proof stage (within approximately 2-6 weeks). The preferred payment method is by credit card; however, other payment options are available

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

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

Sincerely,

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

Associate Editor:
Board Member
Comments to Author:
Thank you for this thorough consideration of our further comments. Great to see the inclusion of a species analysis: another dimension in this really interesting, cross-contextual story.

Appendix A

Review for Proc B: Attending together facilitates social closeness

The authors have carried out an elegant set of experiments that provide evidence in support of idea that bonobos and chimpanzees approach or sit near an individual after they watch something together. The experiments required no training and relied purely on spontaneous response with humans and conspecifics in the case of the chimpanzees. They found the effect in multiple captive populations and were able to put together a nice sample for a relatively powerful test (or this type of work). Below are my main concerns regarding the manuscript. I hope these comments help the authors revise what will be a well cited paper regardless of where it is published.

Conceptually the authors will need to spell things out a bit more for the Proc B audience that is mainly evolutionary biologists. The idea that grooming and play are not shared experiences or that singing and dance is unique to humans will flummox most in this audience. Perhaps for good reason – why should visual contact be prioritized over other perceptual domains and many animals have mating songs and dances that are highly synchronized and can involve intense eye contact (e.g. what is more intimate than the dance and song of the albatross?). To ignore this might work in a psych journal but not here. You will need to explain why vision is so important and animal dancing is different from human dancing and precisely what you mean by "shared experience". Clearly animals share experiences – in that they remember doing things together with others, synchronize and coordinate w/ lifelong partners that they raise young, protect, take risks for and are socially very bonded. This is not what you mean, but you never explicitly state and without you are likely perceived as naïve about the natural world beyond apes. This needs to be very clearly stated. You cannot just reference the shared intentions papers, if u want this to have impact here.

Methodologically I have many minor questions:

Pg 4 ln 10: What was uninterpretable about the 2 chimpanzees you dropped in experiment one? What happened? Can they not be entered into the analysis? Did you try?

Pg 5 ln 10: You note 2 female human strangers were used as experiments. It might be worth citing relevant literature that neither bonobo nor chimpanzee have xenophobic response to strange humans while the two species respond in opposite ways to strange conspecifics. Biologist will be confused otherwise.

Page 6 ln 11: You need to explain the logic of using the baby video. Why was it chosen? How long is it? Did it run on loop? Most important why was only a single video used as the stimulus in both experiments? Does this mean that this effect may only appear when bonobos and chimpanzees watch babies together? Or is it when they watch anything? I think we are left no knowing since you only used one exemplar. Worth noted I imagine.

Page 8 ln 19: You only coded 11 of 48 trials. I would have recommended coded far more than this with a data set this small. You are underpowered here even if the normal guideline of 20% is followed. You also do not state if the coders were blind to condition or hypothesis that I saw.

Figure 2: You have combined the bonobo and chimpanzee data. There are many reasons to suspect the two species might differ in the manipulation and measures used here. You have not justified the combination of the two completely different species data. Typically for a study like this you would provide a supplemental table with subject by subject data that also includes demographic, rearing, experimental history information. Did I miss this somewhere? If so apologies. If not, it needs to be provided.

Why did you not run a 3rd condition with either no TV or a TV with no video running? If both subject were both oriented toward the TV and no video was running wouldn't you just get the same effect? Really this has nothing to do with the shared experience of watching the video? It is just a product of looking in the same direction for some period of time – similar to how direct eye contact increase oxytocin levels and neural synchronization and thus subsequent physical contact….

Page 10 line 15 – 18: is very unclearly stated. It needs to be unpacked. I had to reread several times to determine you had repaired subjects so that they experienced the Exp Con condition up to 4 times. But as stated and with information provided I don't know how many subjects were retested, how many partners they had, who they were partnered with and why. You will need a supplemental table that outlines this subject by subject as is normal here. You should also include demographic information, rearing and relevant experimental history.

In your method sections you do not explain your analytic strategy. This is particularly relevant where subjects were retested and we encounter issues of independence potentially.

Page 10 line 12: I would recommend editing "were unlikely to start fighting" to "are likely to be tolerant of one another"

Page 14 line 10 / 12 you need to operationalize "grooming" and "fighting". Also what safeguards did you employ to protect the safety of your subjects in the case that a fight was out of hand? Were any injured? How was an injury handled? Who judged that and responded? Some sensitivity here is called for.

In experiment 2 I do not see where you report your reliability measurements. Did I miss? Apologies if so.

Figure 4 would recommend labeling Y axis not using title above figure

Page 15 ln 10-11: Odd that the paper is about "closeness" yet all the actual social behaviors that typically represent social closeness remain infrequent. Touching and sitting in contact are typically signs of social closeness in animals like primates yet your manipulation does not alter

these rates in a significant way. While I appreciate the presentation of the nonsignificant findings that seem consistent and supplemental with what the authors frame as their main finding – Im not sure this data helps or hurts your case. It does not appear conservative on your part you have not considered both possibilities here.

Table 1: The authors make the case that the physical set up with different size adjacent rooms with 3D space should not distract from the fact that subjects were closer in the EXP than CON in study 2 – however I do think for the chimpanzees the wall between them is extremely meaningful. It makes them feel safer to be separated that way. Again your manipulation had no effect. Closeness would have been climbing in the same hammock together for instance not keeping a wall between each other. Are you sure you are really thinking about this from a chimpcentric view point or just a theoretical one?

Page 17 line 8: “are already present in other apes” needs to be edited to “is present in humans through shared descent with other apes” otherwise it reads as if you think humans evolved from living apes.

Appendix B

Dear Editor,

We would like to thank you and the two reviewers for your constructive review of our paper. It has helped us to make the paper more clear, and more accessible to your broad readership. Below we have separated out all the individual comments and provided our reply.

Sincerely,

Wouter Wolf

Michael Tomasello

Editor Comments

More detail on the methods is required in order to fully evaluate the study. For example on the choice of video stimulus,

We have provided more information about the stimulus videos by including one of the videos as supplementary material and adding the following paragraph to the methods section of study 1.

"After E1 had sat down, EC started one of several one-minute videos (see SOM video 1 for an example). These videos were excerpts from a longer video of a playing juvenile chimpanzee. This video was chosen based on a recommendation of researchers conducting studies with chimpanzees and bonobos using video stimuli. The key consideration was that the videos should be (1) interesting enough for participants to sit down and attend to the video and (2) not so arousing that they would elicit stress and/or behavior that would cause them to disengage from the video. We were therefore advised to show participants a video of great apes that were not in their own group, which would certainly capture their interest. However, showing them a video of adult males and/or females might contain cues of dominance or mating, which would make the video too arousing. We therefore decided that a video of a playing juvenile chimpanzee would serve the purposes of the current study best. "

definition of closeness in video analyses

We agree that we need to be more specific about this. The relevant section in the manuscript now reads:
"First we coded the time participants spent grooming (i.e. sifting through the hair of another individual, see; 12), a common indicator of social affiliation in great apes (12,14). Additionally, we coded the time participants spent fighting (i.e. aggressively chasing or using physical aggression) with each other during the three minutes after the manipulation as a counter-indicative measure of social closeness."

and reasons for the absence of control conditions (baseline observations or both looking at "no video").

This is an issue that we extensively discussed when designing the current studies, and we therefore understand that these questions emerged during the review process.

First, we want to make clear that the watching alone condition was designed as a control condition. Given that the main question of these two studies concerns the effect of shared/joint experiences on subsequent social closeness, we wanted to design a control condition that engaged participants in the exact same experience as in the experimental condition, except for the experience now not being shared by an experimenter/conspecific. We have made this clearer in the manuscript by replacing the 'Watching Alone condition' with 'Control condition'.

Reviewer 1 proposes a 'baseline' condition, to control for the previous relationship between the subject and the experimenter/conspecific. In study 1 with a human interactant, there was no previous relationship between the participant and the experimenter; the experimenter was a human the ape had never before seen (explained in the text). In Study 2, the two conspecifics did have a previous relationship. However, the question then is what an additional control condition in which two individuals are watching 'no video' would control for that our current control condition does not. The only difference between this proposed baseline condition and our current control condition would be the absence of a video. As such, this baseline would tell us something about the influence of the video (versus no video), but not about the effect of something being watched together. Furthermore, a baseline condition without a video raises the question (1) what the subjects were watching/doing instead, and (2) if they were watching/doing this together or not (they might end up attending to the wall or the heater together, or maybe each other). Thus, even if we would find a difference between this proposed baseline condition and any of the other two conditions, we would not know what might have caused that difference.

As a non-great ape specialist, I fully agree with referee 2's comments on framing the study with the broader, proceedings B audience in mind.

We agree that the previous introduction might have been written too much with a psychology rather than a biology audience in mind. We have therefore added in a new paragraph in the introduction (the second paragraph) to accommodate a biological audience more. In line with reviewer 2's suggestions, this paragraph includes the discussion of some forms of behavior coordination in non-human/ape animal species, together with an explanation of why this type of behavior (and underlying psychology) seems to be different for humans and other animals. This results in a more clearly formulated research question about the psychological effect we demonstrate in the current paper. The added paragraph states:

"Throughout the animal kingdom, the individuals of many species act in coordination with conspecifics. For example, dolphins often behave in synchrony (12), many bird species coordinate their song and dance in a mating context (13,14), and great apes travel together (15) and sometimes hunt monkeys together (16). But do behavioral interactions in which to individuals

focus on an external stimulus together create stronger social relationships or bonds between participants? To our knowledge, there are no studies examining such a relationship in any nonhuman species, and indeed some theorists have suggested that this method of social bonding might be uniquely human (5,10).”

In addition, we have also made small changes to the abstract and discussion to make sure that the example behaviors mentioned are now more specific to humans.

I was surprised that bonobo and chimpanzee data were not analyzed separately in experiment 1 – a point also raised by referee 2.

Given that there were only seven bonobos in the samples, any statistical analysis on to detect a difference between species would have been underpowered. However, we do agree with reviewer 2 that it is useful to have an overview of the individual responses of the subjects in Study 1. We therefore included an excel sheet with the data for each participant in the supplementary materials, as well as a figure in the supplementary materials showing the response times for each subject per condition, separated by species.

As a non-specialist, it’s difficult to tell whether this, or smaller decisions (e.g. observing for < 3 minutes or actively attracting animals in the approach test), are standard or should be justified.

In experimental research with human children and apes where a measure is constituted by the willingness to engage in certain behavior, providing increasingly salient prompts to get participants to engage in that behavior is standard procedure. There are for example many studies on helping and empathy where helping behavior is elicited through increasingly salient prompts, ranging from a display of a need (e.g. reaching for an object) in the first minute, to a verbal request for a specific object in the last minute.

The rationale behind the three minute observation time in Study 2 was that the dependent measure in this study needed more time, as there were no prompts. However, as it was unclear how long it would take for the effect of the manipulation would wear off, we decided to measure the subjects behavior twice as long as in Study 1 (i.e. 3 minutes instead of 1.5).

I would recommend a brief comment where appropriate on the training required to use juice and the methods used to move individuals into the experimental arena.

We have a footnote to the methods section of Study 1 to clarify both of these things, stating:

“All participants were used to being moved around by keepers. Sometimes they were motivated by placing some food in the area where they had to wait for the study to begin. Furthermore, all subjects in Study 1 had extensive experience with juice tubes and therefore required no training in using them.”

For Study 2, the juice tube training was part of the individual exposure trials. We have added a sentence to make this clearer, stating that:

"*All chimpanzees almost instantly understood the mechanics of the juice tube, and drank from them during the exposure trials.*"

In Study 2, we also added a footnote with regard to moving the animals to the right area, stating:

"*As in Study 1, subjects in this group were used to being moved around by keepers, who sometimes used food to direct them towards the desired location where they waited to be let into the testing enclosure.*"

A statement on statistical methods for each experiment would help keep track of sample sizes.

We have added a description of the statistical test, a statement about normality and the sample size (trials and subjects) in a sentence before each analysis.

Reviewers' Comments to Author:

Referee: 1

Comments to the Author(s)

The authors examined if apes affiliate more with a person or conspecific who had watched a video together. I really like their idea, and the MS is written well.
My major concern is the lack of baseline in each measure (latency, proximity). Thus, the causality of effects is unclear. That is, unlike what they discussed, if the subject was already a friend of the experimenter/conspecific, she might have approached faster and spent time together, as well as, watched a video together with the experimenter/conspecific. I recommend the authors to include baseline measures to exclude this possibility. I appreciated that they tried to control a friendship among the two subjects in exp 2, but more stringent or quantitative control (i.e. baseline measures) may be necessary for this purpose.

See our response to the editor's third comment

Minor,
The title "visually attending to things together" implies that the content of the things does not matter. But there is no data supporting this. Videos instead of things?

We agree and changed the title of the manuscript.

Page 4. Please describe what was "uninterpretable behaviour".

We have clarified this, by adding in a footnote stating that:

"*The videos showed that in these trials the participant did not enter the area in front of the experimenter, but merely walked by. However, the experimenter had interpreted that behavior as an approach and had therefore stopped the trial, meaning that no definitive approach latency could be coded for these trials.*"

Page 4. Please describe what was "physiology interfering with the eye tracking set up"

We have clarified this in a footnote, which states:

"One chimpanzee (Jeudi) had a condition in one of her eyes which caused the eye tracker not to be able to track that eye. The other chimpanzee (Riet) had a sizable swelling in the right part of her upper lip which caused her to tilt her head to the side while drinking from the juicetube. This caused the eye tracker not to be able to track her right eye."

Page 6. Please describe the rationale for using a video of a playing baby chimpanzees, and describe the content in more details (a figure shows a baby was playing with two adults in the wild?). I don't think the following is what you intended, but this reminded me of that in SNS, humans enjoy watching "cute" videos together. But apes are unlikely to do the same. From the observational studies, it probably makes more sense to show outgroup male chimpanzees to male chimpanzees to make their ingroup bonds stronger.

See our response to the editor's first comment.

Page 8. I recall that I was recommended by someone to use ICC rather than pearson's r to measure inter-coder reliability because pearson r can be applied to linear data. Anyway, it is not clear if you use ICC or pearson's R.

Both studies reported an interclass correlation (ICC).

Page 9. Please use examples for "a particular way that apes in zoos relates to humans based on their extensive experience with them"

We agree that this could be clarified a bit more and therefore added the sentence:

"*That is, their elaborate interactions with humans, including frequent testing, as well as their dependence on humans for food makes them specifically attuned to human behavior, and, perhaps, therefore also what humans are looking at*."

Page 12. Same as above. Please describe the rationale for using non-chimpanzee animal video as a stimulus. And please describe in more details the content. The video content may matter for the results.

In Study 2, the non-chimpanzee animal video was only used in the exposure trials, which were aimed to familiarize the subjects with the video set-up. We have added a rationale for this video by stating in the manuscript:

"To reduce the likelihood of carry-over effects from the exposure trials into the experimental trials, the stimulus for the exposure trials was an unrelated video of (non-chimpanzee) animal behavior used in a previous study with human children."

For the experimental trials of Study 2, however, we used the same videos as we used in Study 1. It is therefore unlikely that the non-chimpanzee animal video, which was only displayed in the exposure trials, will have influenced the results.

Reviewer 2:
The authors have carried out an elegant set of experiments that provide evidence in support of idea that bonobos and chimpanzees approach or sit near an individual after they watch something together. The experiments required no training and relied purely on spontaneous response with humans and conspecifics in the case of the chimpanzees. They found the effect in multiple captive populations and were able to put together a nice sample for a relatively powerful test (or this type of work). Below are my main concerns regarding the manuscript. I hope these comments help the authors revise what will be a well cited paper regardless of where it is published.

Conceptually the authors will need to spell things out a bit more for the Proc B audience that is mainly evolutionary biologists. The idea that grooming and play are not shared experiences or that singing and dance is unique to humans will flummox most in this audience. Perhaps for good reason – why should visual contact be prioritized over other perceptual domains and many animals have mating songs and dances that are highly synchronized and can involve intense eye contact (e.g. what is more intimate than the dance and song of the albatross?). To ignore this might work in a psych journal but not here. You will need to explain why vision is so important and animal dancing is different from human dancing and precisely what you mean by "shared experience". Clearly animals share experiences – in that they remember doing things together with others, synchronize and coordinate w/ lifelong partners that they raise young, protect, take risks for and are socially very bonded. This is not what you mean, but you never explicitly state and without you are likely perceived as naïve about the natural world beyond apes. This needs to be very clearly stated. You cannot just reference the shared intentions papers, if u want this to have impact here.

See our response to the editor's fourth comment.

Methodologically I have many minor questions:
Pg 4 ln 10: What was uninterpretable about the 2 chimpanzees you dropped in experiment one? What happened? Can they not be entered into the analysis? Did you try?

See our response to referee 1's third comment.

Pg 5 ln 10: You note 2 female human strangers were used as experiments. It might be worth citing relevant literature that neither bonobo nor chimpanzee have xenophobic response to strange humans while the two species respond in opposite ways to strange conspecifics. Biologist will be confused otherwise.

We have added a footnote stating that *"The chimpanzees and bonobos participating in the study have extensive research experience, and are therefore used to interacting with humans they have not encountered before."*

Page 6 ln 11: You need to explain the logic of using the baby video. Why was it chosen? How long is it? Did it run on loop? Most important why was only a single video used as the stimulus in both experiments? Does this mean that this effect may only appear when bonobos and chimpanzees watch babies together? Or is it when they watch anything? I think we are left no knowing since you only used one exemplar. Worth noted I imagine.

See our response to the editor's first comment.

Page 8 ln 19: You only coded 11 of 48 trials. I would have recommended coded far more than this with a data set this small. You are underpowered here even if the normal guideline of 20% is followed. You also do not state if the coders were blind to condition or hypothesis that I saw.

We had an additional naïve coder code 24 of the 48 trials (50%). The interclass correlation is now $r = .999$ (very high). Also, we added the sentence:
"*The angle of the videos that were coded was such that it was impossible to see what condition that subject was in.*"

Figure 2: You have combined the bonobo and chimpanzee data. There are many reasons to suspect the two species might differ in the manipulation and measures used here. You have not justified the combination of the two completely different species data. Typically for a study like this you would provide a supplemental table with subject by subject data that also includes demographic, rearing, experimental history information. Did I miss this somewhere? If so apologies. If not, it needs to be provided.

See our response to the editor's fifth comment

Why did you not run a 3rd condition with either no TV or a TV with no video running? If both subject were both oriented toward the TV and no video was running wouldn't you just get the same effect? Really this has nothing to do with the shared experience of watching the video? It is just a product of looking in the same direction for some period of time – similar to how direct eye contact increase oxytocin levels and neural synchronization and thus subsequent physical contact….

See our response to the editor's third comment.

Page 10 line 15 – 18: is very unclearly stated. It needs to be unpacked. I had to reread several times to determine you had repaired subjects so that they experienced the Exp Con condition up to 4 times. But as stated and with information provided I don't know how many subjects were retested, how many partners they had, who they were partnered with and why. You will need a supplemental table that outlines this subject by subject as is normal here. You should also include demographic information, rearing and relevant experimental history.

We rewrote those sentences, and have added tables for demographic information for both studies, and the information about the pairs in Study 2. We also added a reference to this table in the manuscript:
"*Supplementary data table 3 shows how many times each individual was tested.*"

In your method sections you do not explain your analytic strategy. This is particularly relevant where subjects were retested and we encounter issues of independence potentially.

We have made it clearer what analyses were conducted (see our response to the editor's eight comment). As in Study 2 the analyses were conducted on the pair level, and each pair was unique, the retesting of individuals in the current within subjects design should not have influenced the results.

Page 10 line 12: I would recommend editing "were unlikely to start fighting" to "are likely to be tolerant of one another"
We agree, and adapted this accordingly in the manuscript.

Page 14 line 10 / 12 you need to operationalize "grooming" and "fighting".

We agree that we need to be more specific about this. The relevant section in the manuscript now reads:
"First we coded the time participants spent grooming (i.e. sifting through the hair of another individual, see; 12), a common indicator of social affiliation in great apes (12,14). Additionally, we coded the time participants spent fighting (i.e. aggressively chasing or using physical aggression) with each other during the three minutes after the manipulation as a counter-indicative measure of social closeness."

Also what safeguards did you employ to protect the safety of your subjects in the case that a fight was out of hand? Were any injured? How was an injury handled? Who judged that and responded? Some sensitivity here is called for.

In order to assure that the reader knows that we made sure no participants were injured, we added the following footnote to the description of the experimental trials in Study 2:
"One of the keepers hid behind the enclosure, and peeked around the corner in case they heard aggressive vocalizations so they could intervene in case one of the animals was at risk for serious physical harm. The keepers never intervened, and no injuries were sustained by any of the participants."

In experiment 2 I do not see where you report your reliability measurements. Did I miss? Apologies if so.

Both studies report an interclass correlation (ICC) as a measure of reliability.

Figure 4 would recommend labeling Y axis not using title above figure

We have changed the figures so that the Y-axis is now labeled, and removed the titles above the two bar graphs.

Page 15 ln 10-11: Odd that the paper is about "closeness" yet all the actual social behaviors that typically represent social closeness remain infrequent. Touching and sitting in contact are typically signs of social closeness in animals like primates yet your manipulation does not alter these rates in a significant way. While I appreciate the presentation of the nonsignificant findings that seem consistent and supplemental with what the authors frame as their main finding – Im not sure this data helps or hurts your case. It does not appear conservative on your part you have not considered both possibilities here.

We understand that behavior such as grooming (or physical contact) is intuitively associated with social closeness in great apes, and it might thus seem surprising that our manipulation did not influence this variable. However, research has shown that chimpanzees only spend about 45 minutes a day grooming one another. As such, by virtue of base rate probability, it is not surprising that grooming occurred so unfrequently in the 3 minutes following the manipulation. We had anticipated that, but it seems that we did not make it clear enough that we perceived the social behaviors as secondary variables in our study (as opposed to proximity, which was our primary dependent variable). We have now made this clearer in the text. Furthermore, we have also made this line of reasoning explicit in the results section of Study 2:

"Furthermore, proximity at arm's length (N = 18), as well as fighting (N = 3) and grooming (N = 1) did not occur frequently enough to do statistical analyses on. With regard to grooming this is not surprising. As research has shown that chimpanzees only spend 6.8% of their daytime grooming (i.e. around 45 minutes a day; 20), a three minute window to measure grooming might simply have been too short for such behavior to occur frequently enough to make statistical inferences. However, as Table 1 shows, the trend of time spent (1) in the same room, (2) at arm's length, (3) grooming, and (4) fighting showed a converging trend."

Table 1: The authors make the case that the physical set up with different size adjacent rooms with 3D space should not distract from the fact that subjects were closer in the EXP than CON in study 2 – however I do think for the chimpanzees the wall between them is extremely meaningful. It makes them feel safer to be separated that way. Again your manipulation had no effect. Closeness would have been climbing in the same hammock together for instance not keeping a wall between each other. Are you sure you are really thinking about this from a chimpcentric view point or just a theoretical one?

We agree that this would have been the case if it were not for two things: First, all pairs/subjects were selected to tolerate each other on a basic level, meaning that seeking safety by separation will not have been that much of an issue for most of the dyads. Second the enclosures were surrounded by bars, and not opaque walls, so that being in another room did not allow the subjects to escape from their conspecific's attention. We understand however, that we need to be more explicit with regard to this point. We therefore added the following sentences to the description of the proximity measures:

"[As such, a proximity measure merely constrained by which room the apes were in did not account for their actual physical proximity in three dimensional space]. Also, it is important to note that the enclosures were surrounded by bars (and not opaque walls). This means that subjects could still see each other while being in different rooms, meaning that subjects could not simply hide away from their conspecific's attention by going into a different room. As such, compared to actual physical proximity, being in a separate room, at least in this context, did not seem to be relevant for measuring social closeness."

Page 17 line 8: "are already present in other apes" needs to be edited to "is present in humans through shared descent with other apes" otherwise it reads as if you think humans evolved from living apes.

We agree that this formulation is clearer and more elegant, and have changed it accordingly in the text.

Appendix C

Dear editor,

we would like to thank you and the reviewers for providing a second, very constructive, review of our paper. It has helped us to make the most of our data and to make the paper accessible for both a psychology and biology audience.

Below we have separated out all the individual comments and provided our reply.

Sincerely,

Wouter Wolf

Michael Tomasello

As referee 1 indicates, this may in part be because the general discussion doesn't offer much in the way of synthesis of the two studies. Nor does it fully place the results in context (just 3 citations), or discuss limitations of the study that you describe in the response. For example, those relating to the control conditions, where relevant citations of studies making similar or different experimental choices are also required. Please revise this section based on the additional comments of referee 1.

We have expanded the discussion of the current version of the manuscript in several ways. We have added a paragraph at the beginning briefly summarizing the results of both studies, making it easier for the reader to understand the overall contribution of the two studies taken together.

"The current results show that great apes behave more socially after an interaction in which they align their attention to an external stimulus. Study 1 showed that both chimpanzees and bonobos approach a human experimenter faster after having watched a video with them, suggesting that this effect can be found in the entire genus Pan. Study 2 replicated these findings in a different sample and extended them by showing that this effect is not limited to great apes' interactions with humans, but also seems to occur in interactions between great apes. As such, the current findings shed new light on great ape social cognition and social behavior, as well as the evolutionary origin of connecting through shared experiences in humans."

The current discussion now also discusses some of the design choices more elaborately, citing studies with humans that have used similar designs and control conditions. This section now also discusses how, based on the current result, it is difficult to compare the effects found in the current study to the effect that eye contact has on great ape social interactions.

"One must, however, be cautious when extrapolating a human shared experience-based bonding mechanism to great apes on the basis of the current studies. The current control condition for both studies was designed to keep all parts of the experience, aside from whether it was shared

or not, as constant as possible, similar to studies on shared experiences in humans (10,11). Based on the current results, it is therefore not possible to know whether the results of the current studies generalize to any stimulus other than the video stimulus we used, or to social activities that do not include watching a video. Additionally, the results of the current study do not give insight into how the effect of sharing an experience compares to other factors influencing social closeness, such as for example eye-contact, which has to been shown to play an important role in chimpanzee interactions (30).”

My main concern remains, which is about the inclusion of bonobos. Sample size is not a good justification for lumping together two species. A rationale should include citations to support an evolutionary/behavioural predication that the two species should respond similarly, a methodological note on the latencies you observed, and citations to show precedent for combining these species in other studies that are not specifically setting out to make a cross-species comparison. You've indicated that the sample size is unfortunately too small for a parallel bonobo analysis. I therefore agree with referee 1 that you should either leave out the bonobos or include species in the model and discuss these results. Indeed, this could be very interesting if you find generality not only across the contexts that you've explored but across species!

We understand this concern. Even though in many comparative studies the two species - the only two members of the genus *Pan* and separated by as little as 2 million years of evolutionary time - are often analyzed together. However, as pointed out by the editor and the reviewer, given the species' different social natures, they might respond differently to the manipulation, and some kind of statistical examination is warranted. We therefore replaced the original analysis with a non-parametric 2 by 2 mixed model analysis, and report the following:

“As the data were not normally distributed and data transformations did not resolve this issue, non-parametric statistics were warranted. To test if the manipulation had an effect, as well as whether this effect was different for the different species, we conducted a two (condition: watching together versus watching alone) by two (species: chimpanzee versus bonobo) non-parametric mixed model with approach latency as the dependent variable (24). Results showed no main effect of species on approach latency (modified ATS (df = 1, 10.90) = 3.98, p = 0.07), nor an interaction effect between condition and species (ATS (df = 1) = 0.01, p = 0.92). Crucially, we did find an effect of condition on approach latency (ATS (df = 1) = 5.61, p = 0.017; Figure 2). Participants from both species approached the experimenter with whom they had watched the video together faster (M = 14.86, SD = 13.18) than the experimenter that had been reading her own clipboard (M = 26.78, SD = 22.93).”

The lack of species x condition interaction shows that the two species did not respond differently to the manipulation. Note that we also decided to change the graphs (for both studies) slightly, making them visually more appealing and providing some additional distributional information.

Finally, thank you for taking care to respond to the referees questions in your response. A note and citations are also required in the text for the child and ape studies upon which you base aspects of your experimental methods and for the sentence "animal behavior used in a previous study with human children".

Although data for that initial part of the child study was collected before the data for Study 2 in the current paper, the paper with that child study (and its subsequent follow up studies) is still under review. We can therefore not provide a definitive reference this paper (yet).

Again, the rationale for 3 minutes instead of 1.5 is weak, and could be considered fishing – some justification from previous papers and/or your own observations would be helpful.

Given that there were no prompts in the current study, and we instead relied on proximity, we decided it was necessary to have a longer time window during which proximity was measured. However, we can see how in the previous version of the paper three minutes (instead of five or ten) might seem arbitrary. We have therefore added a more elaborate explanation in the methods section (under "*Measures*") for why chose to decide on a three minute time window, and not, for example a larger time window. Furthermore, we make it more explicit that these three minutes was all data that was recorded, and that after three minutes the experimenter and the keepers came back to let the subject out of the enclosure. The relevant section now reads:

"As the manipulation in Study 2 was no longer an interaction between an ape participant and an experimenter who stayed in one place, but instead a spontaneous interaction between two apes, we could no longer use approach latency as a proxy for social closeness. Instead, we measured participants' subsequent physical proximity as our main dependent variable, consistent with previous research on social networks in great apes (20). As a secondary dependent measure, we also looked at interactive behaviors after the manipulation. In order to obtain sufficient behavior to analyze in the absence of experimental prompts (as in the first study) we decided that a longer time window during which behavior was recorded was necessary. However, at the same time we also had to consider the possibility that the effect of the manipulation would wear off over time, potentially adding noise to the data. Additionally, the staff and researchers that had worked with these subjects before cautioned us that for some subjects, keeping them in their enclosure with nothing to do while part of their group members had already left for the forest would make them uncomfortable. Based on these methodological and animal welfare considerations, we decided to set the time window during which their behavior was measured to three minutes. After three minutes data collection stopped, as the keepers and experimenter returned to the enclosure to let the subjects out."

"Extensive experience" and "sometimes they were motivated by placing some food in the area" need to be more specific, perhaps with a note in the table per individual/trial.

Unfortunately it would be impossible to give an exact description of how much research experience each subject in the current studies had before being tested, as before the current study

the chimpanzees and bonobos had been tested on a daily basis for more than 20 years. We did however expand the footnote to put this in context for the reader:

"The chimpanzees and bonobos participating in the study have extensive research experience (i.e. testing occurs daily, throughout the year, for over 20 years), and are therefore used to interacting with humans they have not encountered before."

Similarly, the moving around of chimpanzees in Study 2 by the keepers using food is part of the keepers' daily routine, for example when the chimpanzees come back in at night. For the sake of their wellbeing, they try to keep some individuals together and others apart, while still giving everyone enough space. To do so, they use food pallets to motivate certain individuals to come into a specific enclosure. As this is such a daily routine, these logistics are never recorded, making it impossible to report the degree to which different individuals have been exposed to this. However, as in the previous case, the footnote in the current paper is gives more context with regards to these logistics:

"As in Study 1, subjects in this group were used to being moved around by keepers, who sometimes used food to direct them towards the right location where they waited to be let into the testing enclosure. This is very similar to the strategy the keepers use every evening when trying to get chimpanzees in specific sleeping enclosures (to minimize conflict at night) by throwing small food pallets in specific areas. As such, the movement of subjects in Study 2 was in line with the chimpanzees' daily routine."

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)

I appreciate that the authors explained the pros and cons of not including no video condition or baseline. My last comment is to intricate the discussion regarding this issue in the text. For example, the test situation in Study 1 is somewhat reminiscent of the established test in which a chimp has to request for foods from someone facing toward the chimp (attentive) but not from someone facing elsewhere (not attentive). But Study 2 can kill this alternative because the stooge is a chimp.

We have elaborated in the discussion on why we decided on the current control condition, citing studies with humans who have used a similar design in the past (see comments to the action editor above).

In Discussion of Study 1, after "that is, ", I don't really understand what they are really trying to say. Consequently, I feel that the rationale for Study 2 is still weak.

The main point here is that there are certain things that great apes only seem to do when raised in an environment where they frequently interact with humans. To make this more clear in the paper, we have added the example of pointing production and comprehension, which apes do much more readily while interacting with humans than while interacting with conspecifics. This part of the paper now reads:

"These results suggest that visually attending to a stimulus together creates some kind of social connection between a great ape and a human, such that the ape experiences an increased motivation to approach the human experimenter. What is still unknown, however, is whether these results indicate a general psychological mechanism in great apes, or rather a particular way that apes in zoos relate to humans based on their extensive experience with them. Previous research has found that great apes are, for example, more likely to understand and produce an imperative pointing gesture while interacting with humans than when interacting with conspecifics (25–28). Similarly, the elaborate interactions with humans of the sample in Study 1, including frequent testing, as well as their dependence on humans for food might make them specifically attuned to what humans are looking at, and this might not generalize to interactions with conspecifics. In Study 2, therefore, we aimed at replicating these findings but between two chimpanzees (and with a different group of chimpanzees, that is, from the Ngamba Island Chimpanzee Sanctuary)."

For the potential species difference, they can at least compare the latency data statistically and show no significant difference between the species. If they found a species difference, it is better to treat the data separately or include the species factor in LMM and present this additional analysis in SI.

We did now include species as a factor in the analysis of Study 1 (see the third comment for the editor).

In Study 2, as Reviewer 2 said, there may be a possibility for the eye-contact effect. It would be nice if the authors briefly discuss or at least argue against this possibility in General Discussion or somewhere because many readers may wonder the same or similar thing.

The current discussion now elaborates on how it is difficult to compare the effects of sharing an experience with other factors influencing social closeness, including eye contact, citing some developmental work with chimpanzees on the prevalence of eye gaze and triadic interactions in chimpanzee parent-child social dynamics.