Response to reviewers

Dear Dr. Veissier,

Thank you for the thoughtful and constructive comments from you and the two reviewers on our manuscript 'Goats who stare at video screens – assessing behavioural responses of goats towards images of familiar and unfamiliar con- and heterospecifics'. As requested, we have revised the manuscript, taking carefully into account all comments made by you and the two reviewers. Together with this revision note, we have resubmitted a revised version of the manuscript with all changes highlighted.

We hope that the present version of the manuscript has improved significantly and that you might consider this manuscript to be recommended by *PCI Animal Science*.

The material in this manuscript has not been published elsewhere and is not submitted for publication elsewhere. All authors have seen the final manuscript and we all take responsibility for its contents.

Sincerely,

Jana Deutsch, Steve Lebing, Anja Eggert & Christian Nawroth

Goats who stare at video screens – assessing behavioural responses of goats towards images of familiar and unfamiliar con- and heterospecifics

Editor Comments to Author:

Your manuscript is well written and brings results interesting to the scientific community. The manuscript needs however revisions, especially in its introduction and discussion/conclusion. Based on the comments from reviewers and my own reading of the manuscript, I recommend that:

The introduction is shortened (4 pages at present whereas introductions are often no longer than 2 pages) and more focused.

The problem you want to address is stated in the second sentence of the introduction. However, you do not detail the different biases that are encountered with current paradigms and so it is not easy for the reader to understand why the looking-time paradigms allows to eliminate these biases. This should be the core of the introduction.

The rest of the introduction can be easily reduced.

The objective of the study should be more clearly stated: is your objective to know if goats can discriminate between images of human vs conspecific and between images of familiar vs. non familiar being OR to know if a looking time paradigm allows studying socio-cognitive capacities in goats (and you use images of human vs conspecific and between images of familiar vs. non familiar to check this)?

> *Author's response: We have shortened the introduction especially with regard to the theoretical background of individual and social recognition. Our study aims at testing whether a looking time paradigm can be used in dwarf goats to answer questions on their ability to recognise other individuals, in this case whether they are capable of differentiating between familiar and unfamiliar con- and heterospecific faces when being presented as two-dimensional images. Therefore, piloting the looking time paradigm in goats was one primary objective of the study and assessing spontaneous social recognition abilities in ungulates was the second objective. We have now better merged these two objectives, especially in the last part of the introduction (e.g. L.121-124), and also adjusted our hypotheses and predictions to highlight this (see L.126-132, L.140-142).*

You better justify the use of human faces

> *Author's response: We used photographs of humans as heterospecific stimuli, as humans were the only other heterospecific species the goats were highly familiar with. Therefore it was the only opportunity to test for the concept of familiarity in another species than goats.*

The discussion is shortened (6 pages at present which is quite a lot) and focused on answers you provide to the initial problem identified,

Answers might be that goats pay attention to images on a screen (at last looking varies), they react differently to conspecific vs human faces (so such differences in behaviour are promising for researches on discrimination or …), essentially looking time varies (so this seems a good variable to use in further studies using such

paradigms), nevertheless there might be differences in perception that cannot be seen in looking time because it is the same for familiar vs unfamiliar images.

> *Author's response: We shortened the discussion, especially with regard to the first paragraphs and the outlook, and changed it according to our adapted hypotheses and predictions (see. L.451-456, L.465-468, L.473-476, L.536-539).*

The precise reasons why goats don't behave differently in front of conspecific vs human faces or why they behave the same in front of familiar vs unfamiliar faces seems to me secondary.

> *Author's response: We now shortened the discussion of why our subjects did not show differential looking behaviour with respect to the categories of familiarity. Nevertheless, we believe that a nuanced discussion of the results of the looking time paradigm is crucial to show how complex the interpretation of such results can be and that it should be done with caution.*

I am not convinced by the structure of the discussion with headings, especially because they are organised as some grouping of results and not as answers to your objectives.

> *Author's response: We have now changed the structure of the discussion so it better aligns with the questions we wanted to answer in our study.*

You refine the description of methods

> *Author's response: We have amended this in the appropriate places based on the reviewers' comments (see below).*

You address all specific comments from reviewers

> *Author's response: We have responded to all comments from the reviewers (see below).*

Minor comment from my side:

L54: replace "while' by 'whereas' if you want to stress the opposition

> *Author's response: This sentence was meant to be an enumeration not an opposition. We changed the sentence to make it better align with the previous sentence.*

> *L.48-51: "As a result, these paradigms can create participation biases (e.g. subjects may lack motivation to participate in the training) and once-learned contingencies may bias the outcomes of subsequent similar tests (Harlow, 1949; Rivas-Blanco et al., 2023)."*

I invite you to revise your manuscript according to the reviewers and my comments. Please make clear in the manuscript what you change (e.g. highlighting changes in

yellow) and provide me with an accompanying letter where you indicate how you have addressed each comments and where (provide line numbers).

> *Authors' response: Thank you for the constructive feedback! We revised our manuscript according to your and the reviewer's comments and highlighted the changes.*

Goats who stare at video screens – assessing behavioural responses of goats towards images of familiar and unfamiliar con- and heterospecifics

<u>Reviewer Comments:</u>

Reviewer 1

This paper explores how goats discriminate between images of familiar and unfamiliar conspecifics and heterospecifics using a looking time paradigm in a goat-adapted apparatus. The paper is very well written and I enjoyed reading it. While the results deviate from findings in other species, the authors are cautious, considering diverse alternative explanations and acknowledging methodological limitations, offering valuable insights for future goat cognition research.

> *Author's response: Thank you for the positive feedback!*

I have some minor comments that I would like the authors to develop:

Methods:

- L193: Is there a specific rationale for exclusively focusing on females and this particular breed?

> *Author's response: The main rationale for conducting this study with female dwarf goats was due to the restricted availability of male goats at our institute. At the FBN, we house a large dwarf goat breeding herd consisting of female animals and only a small group of dwarf goat rams. Additionally, we tested animals that already participated in a previous study that only used female goats. Nevertheless, it would be indeed interesting to conduct our future studies also with male goats to see potential sex differences.*

- L197: Could prior participation in other experiments have impacted the presented results?

> *Author's response: We cannot preclude that participating in other experiments has influenced the looking behaviour of our subjects. Especially as the subjects participated in an experiment with an automated learning device with photographs being presented on a computer display. Nevertheless, we never observed that our subjects showed the learned response (using the video screen as a touchscreen with their snout to indicate a choice regarding a photograph) from this previous experiment so that it did not seem like they have transferred the task requisites from the previous study to ours. It is key for us to improve our adherence to the 3Rs-principle in animal research and therefore to reduce the total number of animals used for our experiments by enabling repeated participation in different experiments, including those with different experimental designs. As the current study had a piloting character regarding the suitability of the looking time paradigm for future use, testing animals with previous testing background that used similar images seemed like an acceptable trade-off to us. We now added this in the discussion:*

> *L.487-494: "We cannot fully exclude that participating in other experiments might have influenced the behaviour of our subjects - especially as the subjects from our study participated in an experiment with an automated learning device*

> with photographs being presented on a computer display. However, we never observed that our subjects showed the learned response from this previous experiment (using the video screen as a touchscreen with their snout to indicate a choice regarding a photograph) so that it can be considered less likely that our subjects have transferred their learned responses and associated behaviours to our study."

- L282: Could you provide more details on human-goat interactions' qualitative and quantitative aspects? Was this interaction sufficient for goats to be familiar with the experimenters, and was it limited to the experimental period?

> *Author's response: The subjects were habituated to the two experimenters for at least three months before data collection for this study started. In these three months, the experimenters provided the subjects with mainly positive interactions (feeding them with dry pasta, if possible touching and petting them) on a regular basis (once a day, five days per week) which was necessary to work with the animals during the experiment. The experimenters were not animal caretakers per se so that the contact was only initiated when the habituation phase started. We consider this daily and positive interaction as being sufficient for goats to be familiar with the experimenters. We now provide further details on this in the manuscript.*

> *L.263-265: "Familiar humans had almost daily positive interactions with the animals during the habituation phase over at least three months."*

Results:

- L426 and 436: Besides human/goat faces, could any other stimuli (a black dot, another species' face) yield similar results when compared to a white screen?

> *Author's response: We also assessed this in a previous experiment (pilot study in the context of a Bachelor's thesis) where we presented the goats with geometrical shapes (squares and circles) compared to a white screen. Similar to our reported results, the results of this pilot study revealed that the subjects looked significantly longer at the geometrical shapes and also directed their first look more often to the video screen presenting a stimulus. We found no effect regarding a preference for a specific type of stimulus (square or circle).*

Discussion:

- L511: More details on how familiar humans and goats interacted are needed; see previous comments.

> *Author's response: Amended (see response to previous comment).*

- L577: What about the social relationships of the tested animals with stimuli goats on the screen. Do you think a more in-depth study of their relationships could enhance understanding of their looking time?

> *Author's response: The animals that were presented as "familiar goats" on the photographs were all from the same social group as the tested subjects. As a*

result, the stimulus subjects were all of an equal "degree" of familiarity to the tested subjects. In our approach, we did not focus on the social relationships between and preferences of subjects, although these could carry additional information when explaining potential biases or preferences in subjects' looking duration. We will definitely consider this in follow-up studies.

- L703-704: I do not think this is correct, since you have some biases and uncontrolled parameters in your experiment that may have hindered a complete test of the concept.

Author's response: We now specified this in the manuscript.

L.617-620: "The looking time paradigm presented here appears to be generally suitable for testing visual preferences in dwarf goats, while assessing the concept of familiarity may require better controls for confounding factors to disentangle the different motivational factors associated with the presented stimuli."

Reviewer 2

The term stimuli is too vague to introduce a precise theoretical framework. You use two-dimensional visual stimuli, i.e. non modified images of living organisms, from conspecifics and human faces.

Author's response: We now indicate earlier that the stimuli we used were two-dimensional visual images.

L.21-25: In our looking time paradigm, we assessed the attention of 10 female dwarf goats (Capra hircus) towards 2D visual stimuli which were images of familiar and unfamiliar con- and heterospecifics (i.e. goats and humans) using an experimental apparatus containing two video screens.

Human is not any ordinary heterospecific species: I mean that you don't consider human as a predator species? I suppose that a vast literature has been devoted to the fact that domestic animals have a long history with human species across domestication. That human is a familiar species of goats. You could have used another heterospecific species, such as a related capridae, a phylogenetically distant species, a predator… So why using human faces?

Author's response: We used photographs of humans as heterospecific stimuli, as humans were the only other heterospecific species the goats were highly familiar with. Therefore it was the only opportunity to test for the concept of familiarity in another species than goats. Due to the experiences the subjects had with their human caretakers and their long history of domestication we don't consider them to perceive humans as a classical predator.

Indeed the introduction does not clearly posit the theoretical challenge(s). Your hypotheses (and predictions) are largely inspired by several studies investigating these questions using domestic and non-domestics animals.

Goats who stare at video screens – assessing behavioural responses of goats towards images of familiar and unfamiliar con- and heterospecifics

> *Author's response: One of our primary focuses was to pilot the looking time paradigm in goats – we now emphasise this more in the introduction (see L.121-124). The second primary focus was to assess spontaneous social recognition abilities in ungulates. For this reason, we chose a study rationale that has already been administered in a relatively similar manner in ungulate species (e.g. Coulon et al., 2011), but not yet in goats.*

> *Coulon, M., Baudoin, C., Heyman, Y., & Deputte, B. L. (2011). Cattle discriminate between familiar and unfamiliar conspecifics by using only head visual cues. Animal Cognition, 14, 279-290.*

Another concern regarding your results is the lack of data regarding the look of the empty screen. Thus, it is not easy to assess whether the faces are highly attractive, i.e. do goats spent much more time looking at the non-empty screen than at the empty one. What would have obtained if the goat were submitted to two empty screens?

> *Author's response: In Figure 5, we show that attention is substantially higher for S+ compared to S-. However, we agree that [S+ - S-] might offer a better picture on the relative attention towards certain stimuli classes. We have re-run the LMM on looking duration with [S+ - S-] instead of [S+] and did not find different outcomes in this new model compared to the old one. We opted for staying with the outcome of the old model for this manuscript.*

The discussion is too long.

> *Author's response: We have now shortened the discussion by excluding excessive examples for potential confounding factors, as well as some parts of the methodological outlook.*

Comments regarding the text

L. At which height above the floor were the screen and the cameras recording frontally the goats positioned? I found the information in fig. 2. When looking at the horizon, at which height the gaze of goats hit the screens?

> *Author's response: Subjects standing in front of the apparatus were considered to look approximately at the centre of the screens but the exact height at which the gaze of the goats hit the screens was dependent on their individual size and movements. We now also added additional information in the manuscript.*

> *L.202-205: "The experimental apparatus (Fig. 2) was inserted into the wall between the testing area and the experimenter booth at a height of 36 cm above the floor and consisted of two video screens (0.55 m x 0.33 m) mounted on the rear wall of the apparatus."*

L. 313: when did you decide to start the data collection from the entrance in the testing area?

Goats who stare at video screens – assessing behavioural responses of goats towards images of familiar and unfamiliar con- and heterospecifics

Author's response: The start of the data collection was adjusted individually for each session based on the behaviour of each subject. When the goat entered the testing area, approached the apparatus and was standing relatively calmly in front of it, the session was started. Sometimes, individual subjects first explored the testing area before approaching the apparatus which led to a delay in starting the session.

Data collection: Did the goat always face/look towards one of the screen after eating the food reward.

Author's response: No, there were (very few) trials in which the animals looked at neither video screen. As indicated in L. 382-383, the trials in which the subjects did not look at S+ at all were removed from the data analysis as in an indication that they might have been distracted.

Did you respect a time delay between the presentation of images within a same test trial?

Author's response: As at least one motivational trial alternated with a test trial, the four stimuli of one session were presented with a break between stimuli presentation of at least 10 seconds.

L. 228: the term "back" wall is misleading: it is in the wall separating the two screens.

Author's response: Thank you for flagging this. We rephrased this sentence.

L.206-210: "Two digital cameras were installed: one (AXIS M1135, Axis Communications, Lund, Sweden) on the ceiling provided a top view of the subject, and one (AXIS M1124, Axis Communications, Lund, Sweden) on the wall separating the two video screens provided a frontal view of the subject."

Video coding: how many did different persons code the video?

Author's response: Unfortunately, a line number is missing so that we are not sure to what line this comment refers to. Regarding the video coding of the looking duration we refer to this notion in

L.333-336: "Inter-observer reliability for the looking duration towards S+ was assessed in a previous stimulus presentation study using the same coding rules and was found to be very high (80 out of 200 trials (40 %) of the videos were coded by two observers; Pearson correlation coefficient (r) = 0.96; p < 0.001)."

We now provide additional information for the video coding of the ear positions:

L.357-358: "All videos were coded by one observer."

L. 334: what was the frequency of video frame acquisition?

Author's response: We now provide further details in the manuscript.

L.210: "Videos were recorded with a 30 FPS rate."

L. 339: it is unclear what is "a head movement directed towards a screen": could you use the line joining the middle of the snout and cutting the segment joining the two eyes to decide objectively whether a goat was looking a screen; for instance it seems difficult to score that variable when the gaze is oriented in the middle of the two screens? See also l. 346

Author's response: We have now clarified this. We acknowledge that there might be a few occasions where gaze might be difficult to score, but in most cases this could be with high precision, which is supported by our high inter-observer reliability scores.

L.315-317: "The first look was scored when the subject directed its gaze towards a video screen for the first time in a trial once the head was lifted from the food bowl."

L 343: did you draw this line manually on each of the frames

Author's response: We drew a fictitious line for each frame, using the illustration only as a template.

L 346: unclear what is the wall containing the apparatus

Author's response: Thank you for flagging this. We have now clarified it.

L.322-325: "The goat's looking behaviour was not scored when the subject was not facing the wall of the testing area in which the apparatus was inserted because then it could not be ensured that it was actually paying attention to the presented stimulus."

L. 349: I discover that goats were climbing…what does this mean?

Author's response: We have now clarified this.

L.325-327: "Video elements in which the goat's face was not visible due to occlusion (e.g. when the subject was sniffing a video screen after moving into the apparatus with both forelegs) were not scored."

L. 350: how could you determine the snout was above the eyes' level because it is possible only using a camera positioned laterally

Author's response: The camera was placed at a height where it could only be directly sniffed by tilting the head back so that the snout was above eye level. There was no other physical option due to the small height of the dwarf goats.

L. 352-353: what is the bottom wall: I do not understand this sentence

Author's response: We have now clarified this part.

> *L.330-333: "There was also no scoring when the subject's snout was perpendicular to the bottom of the apparatus, as in this case it was assumed that the subject was sniffing the bottom of the apparatus with its sight also directed towards it rather than towards the video screens."*

L 378: coding ear position is the most contentious coding work: did you estimate the concordance index in this case. I do not understand how you could detect the 4 positions using only a frontal camera…

> *Author's response: Coding the videos frame-by-frame enabled us to put the single frames into the course of ear movements. Using similar coding criteria for ear positions in a follow-up study had very high inter-observer agreement, further supporting the feasibility of our ethogram using a frontal camera.*

L385: why did you use first a non-parametric test and then a linear mixed-effect model (supposed to be more complete)?

> *Author's response: We have considered different options in how to deal with the data collected in each trial regarding for S+ and S-. Using a ratio of both would have led to a bimodal distribution of data points which is very difficult to implement into linear models. We also did not want to dichotomise our data (e.g. scoring all trials where S+ was higher than S- with a 1, otherwise with a 0). Ultimately, we opted for using the current approach as this would at least allow us to account for dependency between the data points for S+ and S- in each trial.*

I am unclear whether the fact that familiar vs non-familiar images (paired in every sessions) effect was incorporated in the model…

> *Author's response: We are unsure whether we understand the inquiry correctly. The images have not been presented as pairs in each trial but against a white video screen in separate, sequential trials during one session. In the manuscript we state that in the linear mixed effects models "Stimulus species" (two levels: human, goat), "Stimulus familiarity" (two levels: familiar, unfamiliar) and "Testing order" (two levels: first human stimuli, first goat stimuli) were implemented as fixed effects (L.398-401).*

L. 393: it is the total looking duration over 10-sec

> *Author's response: We now added this information.*
>
> *L.373-378: "The four respective response variables were "looking duration at S+" (out of the total of 10s of stimuli presentation), "Forward_Ratio" (time ears oriented forward divided by the summed-up durations of all four ear positions), "Backward_Ratio" (time ears oriented backward divided by the summed-up durations of all four ear positions) and "Horizontal_Ratio" (time ears oriented horizontal divided by the summed-up durations of all four ear positions)."*

Goats who stare at video screens – assessing behavioural responses of goats towards images of familiar and unfamiliar con- and heterospecifics

The legend of figure 2.

> *Author's response: We are unsure on what needs to be adapted in the legend of Figure 2.*

Results

Fig 7 cannot be read. Why didn't you use the same representation as in fig 6?

> *Author's response: Figure 7 (now figure 6) presents model estimates for the LMM on "looking duration at S+", while Figure 6 (now figure 5) refers to paired means of looking time data points for each individual. We understand that increasing layout homogeneity is key for reader comprehension, but do not see how we could adjust this given the different statistical approaches.*

Preference for S+ vs S-: is there an effect of trail session on the FL?

> *Author's response: We did not assess this (see our argument for not using a LMM for this comparison above).*

Difference session 4 vs session: did you compare looking duration during session 1 and session 5. Your difference session 4 vs session may represent an habituation and dishabituation effect…

> *Author's response: Assessing a habituation-dishabituation effect was exactly the purpose of this comparison, we now indicate this earlier in the manuscript:*
>
> *L.398-401: "Moreover, it was tested whether there was an increase in the looking duration towards S+ between session 4 and session 5, due to a dishabituation effect in the subjects caused by the switch of the presented stimulus species."*
>
> *We have not looked at general habituation patterns, but will note this for future studies as these might differ dependent on the biological relevance of the stimuli presented.*

Discussion

L. It is unclear for me

> *Author's response: Unfortunately the line number is missing which is why it is not clear to us to which line this comment refers to.*

L. 477: "assessing attention and arousal" was not one of your objectives. Also how could you measure both of these variables?

> *Author's response: We assessed the visual attention by measuring the looking duration towards the stimuli and explored the ear positions during stimulus presentation as potential indicator for differences in the arousal of goats. We have now amended this in this part of the discussion.*

> *L.454-456: "To assess visual attention (via looking time) and arousal (via ear positions), we measured the goats' looking behaviour towards the stimuli and their ear positions during the trial."*

L 491-496: are these results a surprise? If you had presented a tree (instead of a goat/human) or even a predator vs an empty screen, what would have you expected?

> *Author's response: It is adaptive for a prey animal to direct its gaze to a biologically relevant stimulus compared to a white video screen. We would expect the same outcome for other stimuli. Nevertheless, the degree of biological relevance (e.g. predators being highly relevant and trees being of less relevance) might have an influence on the absolute looking duration towards the stimuli. We will compare looking times towards biological groups of different assumed biological relevance in upcoming studies.*

L. 510: what do you mean by predisposition?

> *Author's response: When we used the term "predisposition", we refered to an innate willingness to develop a certain characteristic, such as a preference for conspecific phenotypes. We understand that this can be hard to disentangle from ontogenetic factors, so we rephrased the sentence accordingly.*

> *L.483-484: "This principle should similarly apply to goats, given their highly social nature, either as an inherent trait or influenced by developmental factors."*

In addition to your proposition, you could also discuss the case where goats never in contact with humans would have been confronted to a human face…

> *Author's response: Our study aimed at comparing two stimulus species and two categories of familiarity (familiar/unfamiliar). Discussing the case where goats never confronted with humans would have been presented with human faces would have been beyond our study purpose. Nevertheless, in the (unlikely) case in which goats never in contact with humans would be tested, we would expect that they would show longer looking durations towards the human images as human faces might be perceived as a predator cues – or would at least induce a strong novelty effect. Due to the request to shorten the discussion, we opted to not include this interesting notion.*

L. 526: It is difficult to interpret your results as an avoidance: the total duration of looking towards human face is not nil.

> *Author's response: It is indeed necessary to interpret or results with caution, especially as multiple motivational factors can simultaneously influence the looking behaviour towards the 2D images, including e.g. a stress buffer effect of conspecific faces AND a possible tendency to avoid human faces. Nevertheless, for avoiding a stimulus, goats would still need to have a quick look at the stimulus, otherwise they would not know what is presented on the screen. Therefore, the looking duration would usually not be nil in an avoidance scenario when scored from the onset of the stimulus presentation.*

Goats who stare at video screens – assessing behavioural responses of goats towards images of familiar and unfamiliar con- and heterospecifics

To what extent the fixation of human faces is different from the empty screen.

Author's response: Please see our comments on the use of [S+ - S-] instead of [S+] above.

L 538: you cannot conclude that the difference of time looking a human image (session 4) and a goat image (session 5) represent/is an habituation-deshabituation mechanism. You have the data but you do not exploit them. Thus you may only suggest that the data could correspond to habituation-deshabituation. Because you pretend using a counterbalanced methods, did you observe the same thing with animals first presented to goat images and after human images.

Author's response: We analysed two groups of Nigerian dwarf goats (group A: 6 subjects, group B: 6 subjects) in a true counterbalanced experimental design. Each subject completed eight sessions: 4 consecutive sessions with goat stimuli, and 4 consecutive sessions with human stimuli. Group A was presented with the goat faces first, while group B was faced with the human faces first. The switch of the stimulus happened only once during the entire experiment, between session 4 and session 5. Only these data can be used to analyse the habituation-dishabituation paradigm.

Effect of stimulus familiarity: your methodology does not allow to test whether the goats discriminate familiar vs non-familiar faces. For instance, if you present simultaneously the two kinds of images, perhaps should you find different results.

Author's response: Indeed, the term "discriminate" can be misleading on some context, as it is often used in a setting where individuals have to learn to respond to certain contingencies. We have changed several parts of the manuscript (L.123, L.141, L. 460-461, L. 568) to decrease ambiguity in writing.