**Reviewer(s)' Comments to Author:**  
  
*Referee: 1*  
  
Comments to the Author(s)

*Paragraph 1) This study contributes to the literature on how decisions are made in groups and whether there are leader-follower dynamics. In its current form the analysis needs to be developed much further to be convincing and rigorous, and the sample needs to be expanded to other groups of goats. It also does not seem of sufficient novelty or general interest to be published in Royal Society Proceedings B.*

*Paragraph 2) A major concern is the small sample size, all data and hence conclusions come from only “a group of adult female goats (n=16)”, meaning that essentially the sample size when drawing conclusions on group behaviour is 1; there is no replication. All the conclusions are also based on only “n=21 decisions”, which are those that meet the authors’ criteria for inclusion (thresholds are often used to categorise behaviour from continuous data such as speeds, but there is no sensitivity analysis to demonstrate that the results are insensitive to these chosen thresholds).*

Given that our conclusions are drawn from data with one herd of goats, we have now couched our results as a case study, testing the novel framework/dataset we present (Lines xxx-xxx).

We have undertaken undertaken a sensitivity analyses to determine how different time windows impact the main dependent variables (supplementary material). We find…

*This makes the study a pilot which needs replicating in different herds, rather than something that can be published as it is. One solution without further data collection would to do some serious modelling (more detail below) and make the model the main part of the paper, with this initial data from a single goat herd as a test for the model.*We now present the goat data as a first test of the conceptual model we present (Lines xxx-xxx), and acknowledge that this work needs to be replicated in more herds/species/contexts (Lines xxx-xxx).

Modelling our goat’s social interactions would be subject to the same scrutiny with regards to our sample size. Interaction “rules” pulled from our small dataset may not be representative. More work on this potentially biased “pilot” may not be advantageous. Instead, we have suggested lab systems to be a better starting point for our test. Sensitivity analysis falls under the same bracket, more rigorous exploration of a preliminary (and potentially invalid – see comment below under Paragraph 4) dataset is not a good use of time. Sensitivity analysis has been recommended for future practitioners who – as was the case in our study - cannot control critical decision points.

*Paragraph 3) It is not clear that from the observational data presented that “individuals’ preferred movement direction prior to departure” can be quantified. How do the authors know what the preferred directions are, and when a decision has really been made? The decision can come before any obvious change in behaviour.*

We don’t know what preferred directions are, only the true directions in which goats point. Decisions can indeed come any time in the process. This is what we test, align then move, or align during movement (the former would suggest decision before moving).

*Paragraph 4) Also, how does the method that the authors present distinguish social decisions (copying or voting) from all the goats in the group responding independently to an external factor? So a non-social decision to move. The only comparison considered is copying versus voting, but it is feasible that all individuals can respond to the same external factor in a seemingly synchronous manner without there really being any social influence. This needs to be considered and ruled out, especially in a field study which the authors argue is a strength of the study.*

This is another valid point. Perhaps the goats were all responding to something other than social interactions (and this is not quantified). This is an important discussion point, and as such is now included in the text. I hope our article will open discussion between field practitioners about how to more holistically quantify environmental variables.

*Paragraph 5) Is it not possible that the collective decision is made before the movement actually starts, after all they all need to point in the same direction early on to move without the group splitting up. One way this could happen is if the goats are using a voting mechanism that the authors are not measuring. For example, the decision could be made by calling, as the authors describe for other species in the Introduction, but calls were not recorded and analysed here. The rationale underpinning the study is that body orientation can potentially allow the goats to ‘vote’ (line 74 onwards) and this is the only cue/signal they can use. This is too simplistic and unrefined, relying on too many assumptions that are not justified with evidence, and there is no systematic ruling out of other potential mechanisms. Without this, the interpretation of the data and conclusions are not convincingly robust.*

Before recent work whereby meerkats were shown to follow a “vocal hotspot”, it seemed that body orientation was the only known mechanism non-human animals can use to communicate directional (rather than timing) preferences. This is no longer the case, though in future work, if vocal signals can be excluded, I believe our approach is valid. This is now very clear in the abstract and the main text of the paper.

*Paragraph 6) The approach to address the major aim of the study, distinguishing between voting and copying, appears to be unreliable. It relies on determining whether the lag between the decision parameter and group speed is negative versus marginally negative or positive. But there is no justification for when a negative time lag is negative ‘enough’ to support voting rather than copying. In other words, in Fig 1a, how close does the curve need to be to 0 to conclude they are copying? For one of the main results used to argue against voting, the confidence intervals are large: mean [95% confidence intervals] = -1.1 [-5.1; 3.0] seconds, suggesting the estimate of the time lag is not accurate enough in this (small) data set.*

I agree with the second part of this comment, that the confidence intervals are too large in our small dataset. However, the premise is sound. We are asking whether time-lag and symmetry of the cross correlation between speed and the decision parameter are significantly negative. These are the simple questions that our statistical test is asking. If the confidence intervals both fall below zero, voting mechanism is supported, and if not, copying. Though understandably, we would expect a more convincing study and window of error before making strong conclusions about copying. Leadership interactions (in supplementary) add additional support to copying in our goat system. Though, I will stress, these “interactions” might be seen in a voting group, by chance alone, and are not enough to discount voting (as many other studies have assumed). *Paragraph 7) The approach taken with the empirical data would be much more convincing if the authors produced a spatially explicit individual-based model where individuals were programmed to either copy or vote, giving quantitative predictions as to how their parameters are expected to change between copying and voting mechanisms, instead of the speculative predictions given in Figure 1.*

I believe the reviewer is referring to spatially explicit models with programmed rules of interaction, taken from our data. Agent based models, which have been used extensively in the study of the collective behaviour of fish and bird swarms can be informed by real data. Though, in a group which - more often than not – appear to not respond to each other’s movements (from a timestep-by-timestep, pairwise perspective), modelling a new type of movement would actually be a significant undertaking worthy of publication as a stand-alone piece. The rules, based on my own observations in goats would look something like this: move once in a while, more likely if neighbour is also moving but not necessarily, not more likely to move towards or align with closer neighbours, just not neighbours which are too far away. So it might be possible to calculate and parameterise probabilities of moving in a time based model. Though work has started in another research group on copying mechanisms which do not rely on a strict temporal component. This latter approach may have more validity in terrestrial groups, where individuals have more time to make decisions, than a “probability of moving” approach I suggested above.

As for the comment about the approach being speculative. The approach taken in figure 1 represents a deduction that if body orientations can predict the departure direction during a “voting with body orientation phase”, the whole group’s mean orientation will either partially (in case of a sub-group vote) or fully (in case of a whole group vote) align with the departure direction during such a phase. The assumptions of the claim that animals vote with body orientation in a strictly controlled voting phase can be questioned (not the least by myself who believes the voting mechanism is an unlikely one), though, if they do, this prediction is a mathematical certainty not a speculation.

The circular mean can be divided into two components. The voting subgroup and the rest of the individuals. After a voting phase, the mean of the voting subgroup will point pretty much directly at the departure direction. Now when this happens, the mean of the whole group (voters and non-votors) will also be skewed toward the direction of departure. If the subgroup is small, this will have a small effect. If it is large (or all the individuals) which vote, this mean value may converge on perfect alignment with the departure direction.  *Paragraph 8) The conclusion that the goats are copying one another would be strengthened if the authors, after using the approach justified in Fig 1, also analysed the spatial structure of this copying behaviour. Presumably the goats are matching their orientation to nearer neighbours if they are copying, so this should be something they can demonstrate from the GPS data. This is essential to support the idea of copying, the results are currently minimal and there is a real need for further analysis to support the conclusions. The paper is quite short so there is plenty of space for this.*

I see the potential in disregarding copying movements which cannot be demonstrated to have occurred between individuals in close proximity. After all one could conclude that significant correlations of heading trajectories would be found in goats on the other side of the planet (every so often). This is why we chose the threshold for correlations over 0.95 (greater than the 0.9 threshold used in the original study from which the leadership methods were attained, see in text for reference). We still found 762 leadership events during the departure phase, and 857 in motion. We also performed the analysis at correlation thresholds of 0.4 and 0.9 and found no difference to the outcome or direction of our results. Altogether, testing whether individuals are copying one another, and at which distances could be fruitful, though in our dataset (which was chosen for cohesiveness of the goats), all seem to respond to one another independent of distance. (An interesting finding in itself.) *Paragraph 9) It is odd that the word ‘goat’ doesn’t appear until the end of the summary, instead “free-ranging social ungulates (Capra aegagrus hircus)” are referred to initially and “a social ungulate (Capra aegagrus hircus)” in the title.*

We have added “goat” to the title.

*Referee: 2*  
  
Comments to the Author(s)

*The main finding of this article is that groups of goats become more aligned as they move and that alignment and movement take place roughly at the same time.  
There is no evidence for voting based on alignment used as a signal, which was observed in other animal species (African buffalos, cormorants, primates). The results do not allow to rule out the possibility that goats use other signals than body orientation to communicate.*

I believe this is addressed above. *My major concern is that the analysis focuses on a short period before departure, but already 100 seconds before departure the decision parameter (which I understand indicates the alignment with the final direction) is much higher than random (figure 2). Hence it is possible that goats are actually 'voting' by pointing to the next direction of movement, only they have started doing so well before the actual departure. The fact that the decision parameter further increases around the time of departure is inevitable (it does so by definition, if it is the alignment of the group direction with the direction of movement at time 0 it must be equal to 1 at time 0; see however my point below).*

100 seconds is by no means ideal, though neither was n=21 decisions. This was a result of a trade-off between number of decisions and length of time analysed. Given how fast a group can change trajectory in other systems (i.e. birds, humans), 100 seconds seemed plenty, to test that if voting with body orientation was a general mechanism which groups use to get from A to B, this timeframe should have captured the phenomemon.

Despite this defence, the reviewer has alerted me that the goats may have used a “subgroup vote” mechanism, and the voting phase started before the 100 second time window. I have included this point in detail in the re-write of the discussion.

*I must say that the description of these parameter is rather unclear: at line 122 the decision parameter is described as 'body orientation relative to the final destination heading', which suggests that for each individual an angle is calculated between current heading and the average heading of the group after departure. However, if this were the case it would be extremely unlikely to have such high values (decision parameter = 1 for time = 0) in figure 2b: this would imply perfect alignment of all members of the group.*

My apologies go out to reviewer 2 here. The decision parameter is now more clearly defined in the text, though I will attempt an answer more specific to the point here.

Regardless of how aligned the individuals are the group can have a decision parameter value of 1. Because the value is a comparison of only two values i) the mean orientation of the group and ii) the departure direction. This is (almost) the way it was calculated by Herbert Prins in the classic buffalo study too. The mean orientation (of adult cows during a “voting phase”) was measured against the departure direction.

*The equation is given for the 'decision parameter' (equation 1), but a(t) and n are not explained. My intuitive understanding is that n is the number of group members and a(t) is the orientation of each group member (it does not make any difference if relative to the north or to a future average moving direction of the group as long as the reference direction is the same for all individuals). If this is the case, however, this is just the modulus of the average alignment vector, isn't it? The subsequent lines 263 - 269 describe in words the calculation of a 'polar order parameter', without giving an equation, but from the description in words it seems to me that this is identical to equation 1, except that instead of calculating mean(cos) and mean(sin), the authors used here sum(cos) and sum(sin) and they normalised at the end by dividing by n.*

Again, I must apologise as ‘n’ sounds a lot like the number of individuals in the group. In fact n is just ‘2’, one is the departure direction and one is the group mean heading. Alpha is heading. This is the general equation to calculate angle similarity (sometimes called: length - *r*). Alpha and n are now more explicitly defined in the text.

*As interpreting a(t) as the orientation of individuals relative to the average direction of the group after departure would not make much sense, we must assume that the reference direction for the 'decision parameter' is different for each individual. Is a(t) the heading of each individual relative to the final destination heading of the same individual? This would explain the graphs and the calculations, but I wonder if it would have any biological meaning: it would be similarly high around departure for a group moving in a consensus direction and for a group which is splitting.*

No it is a group-level measure, not a single individual. We hope the above comment and reworking in the text has made this clearer. *It is surprising (figure 2) that speed and 'decision parameter' increase sharply reaching a plateau already 50 seconds before departure! Shouldn't speed increase around the time of departure? I could not find an explanation for this really strange phenomenon in the text. If 50 seconds before departure they are moving (speed has increased) and they are all oriented towards their final movement direction (decision parameter approaching 1), in what sense they have not departed yet?*

In fact I believe they have (departed). I have changed the wording of “pre-departure” to “departure phase” and subsequent period to “travelling phase”. In reference to the sharp increase, this phenomenon is more a result of combining data from 21 decisions, and is not obvious in the individual group trajectories of single decisions.

*Minor issues:  
  
Figure2: the labels are too small.*

Sorted. *Last paragraph of abstract could expand on the results (support for the copying hypothesis more than for the voting hypothesis), while in a sense it goes back to a more general and vague statement about the need for complex decision-making abilities. I don't think that pooling neighbour directions before departure requires more complex decision-making abilities than pooling them during or after departure. Following consistently the same leader might require some form of individual recognition, but I am not sure that this last sentence of the abstract was intended to point to this aspect.*

This is a valid point and probably arguable from both sides. Therefore I have removed this from the abstract, and from the paper, given the limited space we have to communicate our main idea.  *Caption of figure 3. Please specify for clarity the sign convention used: do positive time lag values indicate that speed changes anticipate alignment changes?*

This is now sorted.