

From disorganized equality to efficient hierarchy: how group size drives the evolution of hierarchy in human societies

Cedric Perret, Emma Hart and Simon T. Powers

Article citation details

Proc. R. Soc. B **287**: 20200693.

<http://dx.doi.org/10.1098/rspb.2020.0693>

Review timeline

Original submission: 27 September 2019

1st revised submission: 27 March 2020

2nd revised submission: 3 May 2020

Final acceptance: 4 May 2020

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

Review History

RSPB-2019-2259.R0 (Original submission)

Review form: Reviewer 1

Recommendation

Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Excellent

General interest: Is the paper of sufficient general interest?

Excellent

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

Good

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

Yes

Is it clear?

Yes

Is it adequate?

Yes

Do you have any ethical concerns with this paper?

No

Comments to the Author

The authors describe a model of hierarchy evolution based on scalar stress, one of the initial analytical models in this literature. The model is rather baroque in terms of number of parameters and built-in assumptions, regarding the interaction of group members, economies of scale, population growth, inheritance, inequality in resource distribution, etc. This adds realism but also makes interpretation more difficult. That said, the authors do a good job in highlighting particular results and the dependency of these results on change in relevant parameters.

I think the model will be of broad interest and favor publication. The model makes a number of predictions that have potential to drive much future research. In fact, the discussion would benefit from more attention to specific predictions the model makes (e.g. how # of leaders interacts with leader-follower influence gap, or how time constraints affect hierarchy evolution) and how these predictions comport with existing ethnographic/experimental data. I also think the authors should in places amend or clarify how they frame the model and how they discuss its implications (see comments below). In particular, the way the authors model how leaders and followers influence each other in groups (the opinion formation model) is central to their findings that hierarchy reduces scalar stress and that hierarchy can evolve on that basis. Since so much rests on this modeling decision, I think more caveats are needed or attention to how such leader-follower interaction may or may not mirror actual group decision-making. Also, the focus is on humans, but why not consider how the model may inform evolution of hierarchy in other species, particularly for the Proc B audience?

Specific comments:

Title: "disorganized equality or efficient despotism" makes it sound like you're modeling hierarchy as a binary, and "despotism" implies coercion (which you're also not modeling per se). Maybe change wording here.

Line 10: Boehm's focus is on the absence of dominance-based hierarchy, and he de-emphasizes or omits the prestige-based hierarchy that often does influence informal political decision-making in these societies. See Garfield et al. (2019).

Lines 15-18: Framing leaders' preferential access to resources or mating partners as exploitation is misleading, to the extent this results at least in part from voluntary exchange. Furthermore, leaders likely benefited reproductively throughout human evolution, even where there were

minimal material resources to contest and an egalitarian ethos limited the extent of hierarchy. See meta-analysis by von Rueden and Jaeggi (2016).

Lines 22-26: “equality” is misleading. Differences in informal political influence according to prestige or gender or age were likely commonplace. Furthermore, sustaining relative egalitarianism may counter-intuitively depend in part on motivations to acquire status via cooperation, not just anti-dominance behavior by followers. You only describe the latter. When status depends on demonstrating value to others (such as through effective leadership), this can result in transfer of knowledge, resources, or reputation from higher status to lower status individuals, limiting growth in status skew over time. See von Rueden et al. (2019) for an empirical demonstration of this using longitudinal ethnographic data.

Lines 38-40: these claims need citation. There is a behavioral economics and a social psychology literature that addresses effects of leaders on things like time to consensus or steering their group towards a cooperative equilibrium, via “first mover” effects or punishment/reward capability.

Lines 41-44: there can also be a conflict over who adopts leader or follower roles, or lack of motivation to adopt leader role at all given its costs- so introducing inter-individual differences into expected costs and benefits of leading may be necessary. See Gavrillets (2015), which takes inspiration from Olson (1965). The traits you later mention, including personality and body size, can affect these costs and benefits.

Lines 45-56: what is meant by scalar stress? Increasing inability to reach consensus/coordinate? Increasing dyadic conflict/free-riding on collective action? Your model focuses on the former, but define scalar stress here because that helps reader anticipate what your mechanistic model will explain. You might cite here Alberti (2014), who fits a simple logistic model to data from Hutterites on group size and fissioning, which he argues has broad cross-cultural validity as a metric of “scalar stress”. Though it doesn’t analytically model the mechanisms that may affect the relationship between group size and fissioning.

Lines 58-63: the discussion of inter-individual differences affecting emergence of informal hierarchy feels like a non-sequitur here. Inter-individual differences inform leader emergence where hierarchy is institutionalized as well. Indeed, your citations here are of studies based on leader emergence in the context of institutionalized hierarchy. Furthermore, inter-individual differences themselves don’t explain emergence of informal hierarchy. You would need to describe how these differences affect acquisition of reputations for an ability to benefit others (prestige) or harm others (dominance), leading to deference to the prestigious or dominant (i.e. hierarchy)- or, as in your model, becoming a speaker and being influenced during speaking events.

Lines 75-78: again, be careful implying an opposition between egalitarianism and leader/follower behavior, without more qualification. Also, before you present these questions, define what you mean by inequality, informal hierarchy, and leadership. You use these terms in discussing the specification and results of your models, but it’s not always clear how they differ. Hierarchy to many readers will imply differences in access to contested resources, but hierarchy has also been defined in terms of differential influence (as revealed in deference signaling between group members). In your models it appears you use the latter definition, but I don’t think this was clear until later in the paper (e.g. Line 284). In your opinion formation model, the distribution of leadership is tantamount to hierarchy, because you require that leaders (i.e. speakers) always have greater influence. Furthermore, greater influence results in inequality: group decisions are more favorable to leaders (i.e. closer to their preference) than the average group member. However, in the evolutionary model, inequality is distribution of production, which is distinct from, though a function of, group members’ influence and interests. Also, in the evolutionary model, hierarchy (as you define it) can coexist with zero inequality (when $d=0$). In other words, leaders might be instrumental in coordinating a final decision that is close to their interests even where the final decision doesn’t differentially benefit them. Right? If so, does it make sense to

have a group decision that is close to one's interests yet doesn't differentially benefit you? This is possible if leaders experience favoritism in subsequent exchange on account of their effective leadership (i.e. their gain in prestige). This more commonly describes the benefits to leadership in more egalitarian settings.

Line 118: define k

Line 124: Dunbar has more recent work that may be even more relevant to this claim, e.g. Zhou et al. 2015.

Lines 129-131: If the listener is more influential, equation 2 can mean the listener's preference moves even further away from the speaker's preference. For example, $a(u)=.5$, $a(v)=1$, $x(u)=.5$, and $x(v)=1$. Then $x'(v)=1.25$. Implementing a rule that speakers are always more influential than listeners by a minimum value (as you do) circumvents this but (a) seems artificial (listeners who are more influential are arbitrarily made less influential than speakers) and (b) cooks hierarchy into the results even more. Influence is already affecting the probability one's own opinion will ever change (based on equation 1).

Line 138: Why define consensus as below a threshold standard deviation of preferences, rather than complete preference agreement? This suggests you are building compromise into your model, where individuals are willing to follow the average view so long as it's below that threshold. Why should compromise be more likely when opinions vary little as opposed to opinions varying a lot? I'd mention this when describing the model in the introduction.

Lines 150-151: very interesting results- particularly that single leaders (as opposed to multiple hierarchy as you call it) are only more efficient re reducing scalar stress when leader-follower influence gap is medium to high. This seems like a good opportunity to describe the importance of analytical models for generating new predictions- how does your result match experimental or observational data comparing number of leaders, leader influence, and group outcomes? In the discussion, you don't address fit of your specific results (rather than their broad implications) to existing data, nor call for tests of specific predictions arising from your model. I would think though that group size and leader-follower influence gap are rarely independent. Managerial mutualism models of hierarchy are built around idea that increasing group size makes group members more willing to offer greater deference or pay a higher fee to managers, to surmount increasing coordination or collective action problems.

Lines 171-174: I would like to see a direct comparison of effect of group size and effect of sampling size (number of listeners) on time to consensus. The effect of sampling size may be related to the effect of group density in driving hierarchy. For example, a study of Amazonian villages suggests that centralization of leadership (i.e. how skewed influence is across individuals) may be more closely linked to residential density than to group size (Glowacki and von Rueden, 2015). In general, I would think group size only matters insofar as it indicates how often people are interacting and in what numbers and how often differences of opinion arise, something density may capture more directly.

Lines 171-174: Perhaps counter-intuitively, your model indicates that there is less scalar stress (lower regression coefficient) the greater the number of listeners per time point (irrespective of number of leaders). And that the number of leaders matters less for scalar stress the greater the number of listeners. Yet intuitively more listeners would suggest more scalar stress, and that leadership should matter more in limiting scalar stress as number of listeners increases. This is a particularity of the model, in that number of listeners is not a metric of number of people trying to reconcile each others' preferences simultaneously. Would be good to emphasize/reiterate this.

Figure 3 caption: I'd stipulate that its number of leaders and number of listeners per discussion event. A casual reader might think the first box of the figure includes analysis of up to 50 leaders

in a group size of only 10.

Lines 273-275: how is the final decision (x^*) generated? Based on the opinion formation model? If so then realized influence is in a sense how one's initial influence updates, right? Based on how well that initial influence determined the proximity of one's preference to the final decision?

Lines 311-314: Figure 4 suggests that as simulation progresses across generations, the range of influence (and possibly skew of influence) becomes much smaller, contrary to what is claimed and what is shown in Figure 5. This just may be because there are so few individuals with influence above 0.5 after 4000 generations that they aren't visible. Maybe mention this.

Lines 340-342: the relationship between time constraints and evolution of hierarchy is another prediction of the model that could be given more time in the discussion. There is a big literature, for example, on how existential threats like war (which you suggest carries greater time constraints) affect choice of leader. Under such conditions, leaders are often granted more influence and dominance in leaders may produce less blowback. So human leader choice may have evolved as a facultative mechanism, if groups consistently experienced activities with different time constraints and thus variable effects of influence skew on fitness. See ethnographic and experimental references in von Rueden and van Vugt (2015):

“Crow and Omaha war captains exercised authority during warfare that was denied them back in camp (Lowie, 1948), Yanomamo headmen from Venezuela were granted authority during raids that they lacked during other times (Chagnon, 1983), the Nambiquara of Brazil endorsed more authoritarian leadership when under threat of attack by neighboring groups (Price, 1981), and Enga leaders' influence in highland New Guinea was heightened during periods of conflict while during peace leadership was challenged and frequently changed hands (Meggitt, 1967).”
 “preferences for larger or more masculine leaders may be greater when group members experience more conflict (van Vugt & de Cremer, 1999) or are under threat from out-groups (Blaker & Van Vugt, 2014; Halevy, Chou, Cohen, & Livingston, 2012; Laustsen & Petersen, 2015; Little et al., 2007; Spisak et al., 2011; Tiguet et al., 2012). In conjunction with our review of leadership in SSSs, these findings suggest that humans evolved a psychology that trades-off the risk of exploitation by physically dominant leaders with their greater coordination and conflict resolution efficiency in face-to-face interaction with followers (Lukaszewski & von Rueden, 2015; von Rueden et al., 2014). This tradeoff is more pronounced during situations when free-riding and coordination failure are probable or pose an existential threat to the group, such as during periods of heightened intra- or inter-group conflict.”

Lines 351-357: you might comment on role of group selection models in driving hierarchy. For example as discussed in Makowsky and Smaldino (2016).

Lines 358-365: this result likely depends heavily on reproductive skew and fidelity of influence inheritance. In actual human groups, differential reproduction of influential individuals has been demonstrated for societies of all levels of hierarchy (see von Rueden and Jaeggi, 2016) so has this made humans more homogeneous in terms of influence? There may also be frequency-dependent mechanisms maintaining differences in influence. Or mutation is stronger than you model. Or horizontal inheritance matters: the effects on influence from one's social network connections (see this demonstrated in von Rueden et al. 2019).

Lines 383-385: this though depends on precisely how the model links listening group size to time to consensus. If leader-follower interactions happened through other mechanisms, such as simultaneous influence attempts across listeners, this result may be qualitatively different, right? Worth mentioning because it's a central assumption of the model affecting the paper's main conclusions.

Lines 429-430: how was hierarchy modeled differently to make it “institutionalized”, relative to the hierarchy modeled in the current study?

Lines 434-437: another key aspect missing is the collective action problem, i.e. costs from free-riding. Scalar stress might involve not just time to consensus, represented as an opportunity cost in your evolutionary model, but also free-riding or conflicts over resource division. Other models of evolution of leadership focus on these costs, e.g. Hooper et al. 2010.

Lines 450-460: per your call for more detailed study of cost of organization in groups, see work on the Tsimane comparing group size, group density, and conflict frequency to particular traits of leaders (von Rueden et al. 2014) and, most relevant to your model, to skew in political influence (Glowacki and von Rueden 2015). I don't know of other ethnographic data (particularly where hierarchies are not institutionalized) that quantitatively compares intra-group skew in political influence to metrics of scalar stress.

- Alberti, G. (2014). Modeling group size and scalar stress by logistic regression from an archaeological perspective. *PLoS ONE*, 9, e91510.
- Garfield, Z., von Rueden, C., and Hagen, E. (2019). The evolutionary anthropology of political leadership. *The Leadership Quarterly*, 30, 59-80.
- Gavrilets, S. (2015). Collective action problem in heterogeneous groups. *Phil. Trans. Roy. Soc. B*, 370, 20150016.
- Glowacki, L. and von Rueden, C. (2015). Leadership solves collective action problems in small-scale societies. *Phil. Trans. Roy. Soc. B*, 370, 20150014.
- Hooper, P., Kaplan, H., and Boone, J. (2010). A theory of leadership in human cooperative groups. *Journal of Theoretical Biology*.
- Olson M. (1965). *Logic of collective action: public goods and the theory of groups*. Cambridge, MA: Harvard University Press.
- Makowsky M., and Smaldino, P. (2016). The evolution of power and the divergence of cooperative norms. *Journal of Economic Behavior & Organization*, 126, 75-88.
- von Rueden et al. (2014). Leadership in an egalitarian society. *Human Nature*, 25, 538-566.
- von Rueden, C. and van Vugt, M. (2015). Leadership in small-scale societies: some implications for theory, research, and practice. *The Leadership Quarterly*.
- von Rueden, C. and Jaeggi, A. (2016). Men's social status and reproductive success in 33 non-industrial societies: effects of subsistence, marriage system, and reproductive strategy. *Proc. Natl. Acad. Sci. USA*, 113, 10824-10829.
- von Rueden, C., Redhead, D., O'Gorman, R., Kaplan, H., and Gurven, M. (2019). The dynamics of cooperation and social status in a small-scale society. *Proceedings of the Royal Society B*, 286, 20191367.
- Zhou, W., Sornette, D., Hill, R., and Dunbar, R. (2015). Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society B*, 272, 439-444.

Review form: Reviewer 2

Recommendation

Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Good

General interest: Is the paper of sufficient general interest?

Acceptable

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

Good

Is the length of the paper justified?

No

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

N/A

Is it clear?

N/A

Is it adequate?

N/A

Do you have any ethical concerns with this paper?

No

Comments to the Author

In this paper the authors present mathematical models aimed at explaining the evolution of centralized decision-making or hierarchy in humans. This is an important topic because this is one of the most striking trends in human history yet relatively little work has been done in the way of developing formal mathematical models of the processes involved.

First an opinion formation model is developed. In this set-up individuals take turns to influence others in a group to take on their preferred value of a trait. Success in the task is measured as time taken for the group to reach consensus. Social structure is introduced by the fact that different individuals can be given different levels of influence in the model. For simplicity two classes of individuals are used in the analysis – leaders (greater influence in discussions) and followers. The initial results indicate that including hierarchy reduces the time taken for groups to make decisions. Where leaders have a strong influence then having a single leader is better as the group only gets pulled to one value.

The authors then establish an evolutionary version of the model. Pay-offs depend on how long the task takes to complete and the distribution of leaders and followers (leaders can receive a greater pay off than followers) – therefore the idea is that leaders may emerge because they enable groups to reach consensus earlier and thus receive a higher pay off. The “influence” parameter is modelled as property of individuals and is allowed to evolve. The model is explored in an island model with different groups on different patches. An additional parameter (d) is added that affects whether those with more influence do indeed receive a greater benefit.

Apologies if I have got any of that wrong but there are quite a lot of moving parts here and the manuscript is not always laid out in a way that makes it easy to follow. Overall I think the model developed here has some merit and could be published. The results seem to make intuitive sense and follow logically from the description of the model. I would like to see some steps taken to improve its clarity and also allow readers to assess its applicability to the phenomena it is interested in explaining. I want to stress that I don't think the authors have done a bad job of writing the paper it is more that the complexity of what they have done makes it challenging to take in and more could be done to break things up and help the reader.

One thing that is challenging is that terms are not always clearly defined or are introduced suddenly without being set-up sufficiently earlier in the paper. For example I am still not entirely sure what the difference the authors make here between informal hierarchy and institutional hierarchy and whether this matters for the model.

For example lines 311-316 state "The results shows that despite the wide range of possible distribution of influence, individuals evolve towards hierarchy, i.e. a minority of leaders with high influence and a majority of followers with low influence. In the meantime, the population grows to a large group size. Within a patch, informal hierarchy also evolves but the proportion of leaders and followers vary."

It is unclear whether "informal hierarchy" within a patch is somehow different to the evolution of hierarchy described in general.

Similarly the authors make a distinction between hierarchy and inequality towards the end of the paper without really making it clear that this was an important point earlier in the paper. Calling "d" the level of ecological inequality is a bit confusing, and introducing ecological variation across patches as well as all the other things at the same time makes it difficult for the reader to understand the importance of any particular process.

It seems at the beginning of the paper that "scalar stress" is a key thing about the model being created yet the definition is very brief "The "iron law of oligarchy" proposes that a key element is scalar stress, which describes the fact that the cost of organisation increases with group size."

(line 45) It would be good to give a fuller description and highlight with a specific example.

Related to this on line 377 or a result about scalar stress is introduced as a secondary finding and this was a little surprising as it seemed that the paper had been discussing this in terms of the effect of leaders in reducing time to consensus – so I am now a bit puzzled by what scalar stress actually is.

In the evolutionary model I was unclear how alpha as an individual trait is used to assign influence in the group (presumably individuals in a group are ranked according to their alpha values and then weight of influence is assigned somehow – but it wasn't clear to me from the text how this was done) – I then found the explanation earlier in the paper – however at this point a simple ascription of followers and leaders had been used so it seemed to make less sense to have that explanation there. The role of listener was also a little unclear at first.

To improve the paper I have two main recommendations; 1) include a figure and/or schematic diagrams that illustrate the model clearly and help orient the reader, 2) Break up the paper into more manageable chunks – use sub-headings. Build up the results by showing the effects of new aspects of the model (e.g. can you show the model results without the "ecological inequality" part first and then add to show the effect it has?).

There are also some modelling decisions that could be discussed further:

"The trait α carried by individuals is now an evolving trait and is transmitted vertically from parent to offspring, e.g. by social learning, as is common in both hunter-gatherer groups [28] and modern societies [29]. (lines 192-195)

The trait seems to be inherited only by vertical inheritance and is in effect no different from a genetic trait – except with a larger mutation rate (see paper by Strimling and McElreath in Current Anthropology). Would expect any differences if you had true cultural transmission? More could be said about the mechanism by which hierarchy spreads between populations and what selective processes are at play here (is there a balance between group and individual selection? This is implied but would be good to make more overt.) To what extent does the modelled process match what goes on in the real world?

The only benefit leaders have in this model is their ability to get individuals in a group to coordinate quickly i.e. there is no difference in outcomes or pay-offs for the particular value the group settles upon. What situations in the real-world does this apply to? What differences might we see if there were different outcomes for different values of the trait?

Related to the last point the applicability and scope of the model would be easier for the reader to understand what situations the model most closely resembles. Rather than being a general model

of the evolution or importance of hierarchy it sounds like the authors are modelling something like the emergence of “Big Men” style societies where certain charismatic individuals attempt to influence members of their group. In the literature this often referred to as “achieved” rather than “ascribed” leadership because there is no office of leader that must be filled and the role is not passed on in an hereditary manner. To fit with the broad scope of PRSB it might be worth saying to what extent this model could apply to non-human animals too and provide examples.

Minor points

Lines 13-14 The independent emergence of hierarchy and its pervasive presence suggest an evolutionary origin [4].

Unclear what you mean by evolutionary origin here

Lines 19-22 “Humans have inherited traits and preferences towards hierarchy from their primate ancestors, who were organised in dominance hierarchies where individuals physically compete for rank, resources and partners [2]”

But this runs counter to the egalitarian ethos that hunter-gatherers are well-known for and even western people in lab studies have been shown to have inequity aversion

Review form: Reviewer 3

Recommendation

Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Excellent

General interest: Is the paper of sufficient general interest?

Good

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

Excellent

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

N/A

Is it clear?

N/A

Is it adequate?

N/A

Do you have any ethical concerns with this paper?

No

Comments to the Author

An impressive paper that joins an opinion formation model permitting influential leaders to emerge, thus limiting degree of scalar stress (promoting efficiency of decision-making) with an evolutionary model in which social organization is cleverly described along a continuum of skew in degree of influence, thus linking the two models. The bottom line is that the organizational advantages of leadership sufficiently limit opportunity costs and permit increasing returns to scale to permit leadership/followership to emerge even when followers are disadvantaged by lower reproduction. Leadership also allows group sizes to grow, allowing returns to scale to be realized.

General comments: this is an exceedingly complex model and anything the authors can do to help the reader fit it all together would be appreciated. For example, it might be helpful to have a single table where all the parameters are defined. The potential parameter space is very large and is only lightly sampled. (I could not see the supplementary materials however as I explained to the editor.)

I had trouble interpreting several of the key figures. On 1 & 2, what is n_{sub1} ? The population structure has not (yet at least) been explained. One assumes that n_{sub1} is the size of the first group, but how does that relate to the x axis where group size ranges from 0 to 1000? Similar issue with 2; how can groups of 30 have 50 leaders? Also 2, I presume that the linear regression coefficient depicted is the slope (should be explicit).

On 4, I wonder if it might be mislabeled. Is it the mean of alpha or its mean skewness that is indexed by the red/blue scale? Does Mean size by patch really indicate Mean group size by patch? What is meant by "Proportion of individuals in the population"? I think what we're supposed to see is that through time skewness of alpha goes up, as does mean group size, but I can't decode it.

On Figure 5C, blue, are those resources per capita, or simply in total? It seems to me that it's the per capita figures that would be critical.

General comment: since leadership has somewhat different effects here (decreasing opportunity costs as group sizes grow, thus allowing group sizes to grow which in turn allows them to enjoy increasing returns to scale) its hard to parse at any point which of these effects dominates. Most likely they are co-constitutive and equally important, but I couldn't glean that from the ms.

Starting in 287: "On the one hand, an efficient hierarchical organisation allows a group to produce a larger amount of resources and hence reach a larger size. On the other hand, hierarchy provides a stronger advantage as group size increases because the cost of organisation increases less in hierarchical groups than in egalitarian groups." What (in the model) if anything ever puts the brakes on this process? What (in the reference societies) might do this?

"Economies of scale" are mentioned a number of times, e.g., "We make the assumption of economy of scale in which additional participants increase the benefit superlinearly [31]." I think the more common usage is to speak of increasing returns to scale, and I think that's what's actually meant (given the reference to superlinear).

There are some missed opportunities to link to other relevant work. For example, in a very recent issue of *Antiquity*, Amy Bogaard and co-authors show why agriculture where returns are limited by labor is less likely to produce wealth inequality than is agriculture limited by land. (I'm traveling and don't have access to the exact citation.) This seems (superficially at least) to be at variance with the statement "The benefit of hierarchy depends on the time constraints C_t , which translates the time to reach consensus into an opportunity cost of organisation. ... For a low level

of time constraints, the benefit of hierarchy has a negligible effect on organisation and group production and thus, hierarchy does not evolve." Why this disjunction? Interesting since both of these statements ostensibly target Neolithic societies.

Are there any empirical data available that could inform us as to whether the levels of alpha used in the model for informal hierarchy are realistic? If so, please cite.

Also, quite a few uncited works come to generically similar conclusions about group size and leadership. Even though these do not (so far as I can recall) contain opinion formation models, it may be that some of these suggest that such a model is not strictly necessary? For example: Crabtree, Stefani A., R. Kyle Bocinsky, Paul L. Hooper, Susan C. Ryan, and Timothy A. Kohler 2017 How to Make a Polity (in the central Mesa Verde region). *American Antiquity* 82(1):71-95. (See also references therein.)

Also, you cite a working paper that has now been published, as follows:

Hooper, Paul, Eric Smith, Henry Wright, Timothy Kohler, and Hillard Kaplan 2018 Ecological and Social Dynamics of Territoriality and Hierarchy Formation. In *The Emergence of Pre-Modern States: New Perspectives on the Development of Complex Societies*, edited by Jeremy Sabloff and Paula L. W. Sabloff, pp. 105-132. Santa Fe Institute Press, Santa Fe.

Trivia: typos of some sort in lines 107, 444; Ref 12 incomplete.

Decision letter (RSPB-2019-2259.R0)

25-Nov-2019

Dear Mr Perret:

I am writing to inform you that your manuscript RSPB-2019-2259 entitled "Disorganised equality or efficient despotism: How group size drives the evolution of hierarchy in human societies" has, in its current form, been rejected for publication in *Proceedings B*.

This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

- 1) A 'response to referees' document including details of how you have responded to the comments, and the adjustments you have made.
- 2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
- 3) Line numbers in your main document.

To upload a resubmitted manuscript, 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 Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Sincerely,
 Professor Gary Carvalho
 mailto: proceedingsb@royalsociety.org

Associate Editor
 Board Member: 1

Comments to Author:

Many thanks for submitting your Proceedings B, and apologies for taking a long time to get back to you with a decision. Your patience is very much appreciated and I hope you feel it was worth the wait. I have received three excellent reviews from experts in fields directly related to the topic of your study, all three of which provide a range of insightful and constructive comments and suggestions. While they all agree that this is an impressive piece of work that provides a number of new and important insights into the evolution of hierarchies, they also all comment on the complexity of the model and express a need for some additional work to ensure your work reaches the broad readership it deserves. I fully agree with this assessment and I would therefore urge you to take this, and all other issues raised, onboard as best as you can. As a matter of fact, I found the manuscript, and the Introduction in particular, very well written. It does a very good job at setting the scene for what is to come, and to introduce the questions at hand to a non-specialist audience. This combined with the fact that you are aiming at explaining an intriguing question that is likely to appeal to many, I think there is the potential for a manuscript that will be of interest to the broad readership of Proceedings B, including those readers that would not necessarily read theoretical papers otherwise.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)

The authors describe a model of hierarchy evolution based on scalar stress, one of the initial analytical models in this literature. The model is rather baroque in terms of number of parameters and built-in assumptions, regarding the interaction of group members, economies of scale, population growth, inheritance, inequality in resource distribution, etc. This adds realism but also makes interpretation more difficult. That said, the authors do a good job in highlighting particular results and the dependency of these results on change in relevant parameters.

I think the model will be of broad interest and favor publication. The model makes a number of predictions that have potential to drive much future research. In fact, the discussion would benefit from more attention to specific predictions the model makes (e.g. how # of leaders interacts with leader-follower influence gap, or how time constraints affect hierarchy evolution) and how these predictions comport with existing ethnographic/experimental data. I also think the authors should in places amend or clarify how they frame the model and how they discuss its implications (see comments below). In particular, the way the authors model how leaders and followers influence each other in groups (the opinion formation model) is central to their findings that hierarchy reduces scalar stress and that hierarchy can evolve on that basis. Since so much rests on this modeling decision, I think more caveats are needed or attention to how such leader-follower interaction may or may not mirror actual group decision-making. Also, the focus is on humans, but why not consider how the model may inform evolution of hierarchy in other species, particularly for the Proc B audience?

Specific comments:

Title: “disorganized equality or efficient despotism” makes it sound like you’re modeling hierarchy as a binary, and “despotism” implies coercion (which you’re also not modeling per se). Maybe change wording here.

Line 10: Boehm’s focus is on the absence of dominance-based hierarchy, and he de-emphasizes or omits the prestige-based hierarchy that often does influence informal political decision-making in these societies. See Garfield et al. (2019).

Lines 15-18: Framing leaders’ preferential access to resources or mating partners as exploitation is misleading, to the extent this results at least in part from voluntary exchange. Furthermore, leaders likely benefited reproductively throughout human evolution, even where there were minimal material resources to contest and an egalitarian ethos limited the extent of hierarchy. See meta-analysis by von Rueden and Jaeggi (2016).

Lines 22-26: “equality” is misleading. Differences in informal political influence according to prestige or gender or age were likely commonplace. Furthermore, sustaining relative egalitarianism may counter-intuitively depend in part on motivations to acquire status via cooperation, not just anti-dominance behavior by followers. You only describe the latter. When status depends on demonstrating value to others (such as through effective leadership), this can result in transfer of knowledge, resources, or reputation from higher status to lower status individuals, limiting growth in status skew over time. See von Rueden et al. (2019) for an empirical demonstration of this using longitudinal ethnographic data.

Lines 38-40: these claims need citation. There is a behavioral economics and a social psychology literature that addresses effects of leaders on things like time to consensus or steering their group towards a cooperative equilibrium, via “first mover” effects or punishment/reward capability.

Lines 41-44: there can also be a conflict over who adopts leader or follower roles, or lack of motivation to adopt leader role at all given its costs- so introducing inter-individual differences into expected costs and benefits of leading may be necessary. See Gavrilets (2015), which takes inspiration from Olson (1965). The traits you later mention, including personality and body size, can affect these costs and benefits.

Lines 45-56: what is meant by scalar stress? Increasing inability to reach consensus/coordinate? Increasing dyadic conflict/free-riding on collective action? Your model focuses on the former, but define scalar stress here because that helps reader anticipate what your mechanistic model will explain. You might cite here Alberti (2014), who fits a simple logistic model to data from Hutterites on group size and fissioning, which he argues has broad cross-cultural validity as a metric of “scalar stress”. Though it doesn’t analytically model the mechanisms that may affect the relationship between group size and fissioning.

Lines 58-63: the discussion of inter-individual differences affecting emergence of informal hierarchy feels like a non-sequitur here. Inter-individual differences inform leader emergence where hierarchy is institutionalized as well. Indeed, your citations here are of studies based on leader emergence in the context of institutionalized hierarchy. Furthermore, inter-individual differences themselves don’t explain emergence of informal hierarchy. You would need to describe how these differences affect acquisition of reputations for an ability to benefit others (prestige) or harm others (dominance), leading to deference to the prestigious or dominant (i.e. hierarchy)- or, as in your model, becoming a speaker and being influenced during speaking events.

Lines 75-78: again, be careful implying an opposition between egalitarianism and leader/follower behavior, without more qualification. Also, before you present these questions, define what you mean by inequality, informal hierarchy, and leadership. You use these terms in discussing the

specification and results of your models, but it's not always clear how they differ. Hierarchy to many readers will imply differences in access to contested resources, but hierarchy has also been defined in terms of differential influence (as revealed in deference signaling between group members). In your models it appears you use the latter definition, but I don't think this was clear until later in the paper (e.g. Line 284). In your opinion formation model, the distribution of leadership is tantamount to hierarchy, because you require that leaders (i.e. speakers) always have greater influence. Furthermore, greater influence results in inequality: group decisions are more favorable to leaders (i.e. closer to their preference) than the average group member. However, in the evolutionary model, inequality is distribution of production, which is distinct from, though a function of, group members' influence and interests. Also, in the evolutionary model, hierarchy (as you define it) can coexist with zero inequality (when $d=0$). In other words, leaders might be instrumental in coordinating a final decision that is close to their interests even where the final decision doesn't differentially benefit them. Right? If so, does it make sense to have a group decision that is close to one's interests yet doesn't differentially benefit you? This is possible if leaders experience favoritism in subsequent exchange on account of their effective leadership (i.e. their gain in prestige). This more commonly describes the benefits to leadership in more egalitarian settings.

Line 118: define k

Line 124: Dunbar has more recent work that may be even more relevant to this claim, e.g. Zhou et al. 2015.

Lines 129-131: If the listener is more influential, equation 2 can mean the listener's preference moves even further away from the speaker's preference. For example, $a(u)=.5$, $a(v)=1$, $x(u)=.5$, and $x(v)=1$. Then $x'(v)=1.25$. Implementing a rule that speakers are always more influential than listeners by a minimum value (as you do) circumvents this but (a) seems artificial (listeners who are more influential are arbitrarily made less influential than speakers) and (b) cooks hierarchy into the results even more. Influence is already affecting the probability one's own opinion will ever change (based on equation 1).

Line 138: Why define consensus as below a threshold standard deviation of preferences, rather than complete preference agreement? This suggests you are building compromise into your model, where individuals are willing to follow the average view so long as it's below that threshold. Why should compromise be more likely when opinions vary little as opposed to opinions varying a lot? I'd mention this when describing the model in the introduction.

Lines 150-151: very interesting results- particularly that single leaders (as opposed to multiple hierarchy as you call it) are only more efficient re reducing scalar stress when leader-follower influence gap is medium to high. This seems like a good opportunity to describe the importance of analytical models for generating new predictions- how does your result match experimental or observational data comparing number of leaders, leader influence, and group outcomes? In the discussion, you don't address fit of your specific results (rather than their broad implications) to existing data, nor call for tests of specific predictions arising from your model. I would think though that group size and leader-follower influence gap are rarely independent. Managerial mutualism models of hierarchy are built around idea that increasing group size makes group members more willing to offer greater deference or pay a higher fee to managers, to surmount increasing coordination or collective action problems.

Lines 171-174: I would like to see a direct comparison of effect of group size and effect of sampling size (number of listeners) on time to consensus. The effect of sampling size may be related to the effect of group density in driving hierarchy. For example, a study of Amazonian villages suggests that centralization of leadership (i.e. how skewed influence is across individuals) may be more closely linked to residential density than to group size (Glowacki and von Rueden, 2015). In general, I would think group size only matters insofar as it indicates how

often people are interacting and in what numbers and how often differences of opinion arise, something density may capture more directly.

Lines 171-174: Perhaps counter-intuitively, your model indicates that there is less scalar stress (lower regression coefficient) the greater the number of listeners per time point (irrespective of number of leaders). And that the number of leaders matters less for scalar stress the greater the number of listeners. Yet intuitively more listeners would suggest more scalar stress, and that leadership should matter more in limiting scalar stress as number of listeners increases. This is a particularity of the model, in that number of listeners is not a metric of number of people trying to reconcile each others' preferences simultaneously. Would be good to emphasize/reiterate this.

Figure 3 caption: I'd stipulate that its number of leaders and number of listeners per discussion event. A casual reader might think the first box of the figure includes analysis of up to 50 leaders in a group size of only 10.

Lines 273-275: how is the final decision (x^*) generated? Based on the opinion formation model? If so then realized influence is in a sense how one's initial influence updates, right? Based on how well that initial influence determined the proximity of one's preference to the final decision?

Lines 311-314: Figure 4 suggests that as simulation progresses across generations, the range of influence (and possibly skew of influence) becomes much smaller, contrary to what is claimed and what is shown in Figure 5. This just may be because there are so few individuals with influence above 0.5 after 4000 generations that they aren't visible. Maybe mention this.

Lines 340-342: the relationship between time constraints and evolution of hierarchy is another prediction of the model that could be given more time in the discussion. There is a big literature, for example, on how existential threats like war (which you suggest carries greater time constraints) affect choice of leader. Under such conditions, leaders are often granted more influence and dominance in leaders may produce less blowback. So human leader choice may have evolved as a facultative mechanism, if groups consistently experienced activities with different time constraints and thus variable effects of influence skew on fitness. See ethnographic and experimental references in von Rueden and van Vugt (2015):

"Crow and Omaha war captains exercised authority during warfare that was denied them back in camp (Lowie, 1948), Yanomamo headmen from Venezuela were granted authority during raids that they lacked during other times (Chagnon, 1983), the Nambiquara of Brazil endorsed more authoritarian leadership when under threat of attack by neighboring groups (Price, 1981), and Enga leaders' influence in highland New Guinea was heightened during periods of conflict while during peace leadership was challenged and frequently changed hands (Meggitt, 1967)." "preferences for larger or more masculine leaders may be greater when group members experience more conflict (van Vugt & de Cremer, 1999) or are under threat from out-groups (Blaker & Van Vugt, 2014; Halevy, Chou, Cohen, & Livingston, 2012; Laustsen & Petersen, 2015; Little et al., 2007; Spisak et al., 2011; Tigue et al., 2012). In conjunction with our review of leadership in SSSs, these findings suggest that humans evolved a psychology that trades-off the risk of exploitation by physically dominant leaders with their greater coordination and conflict resolution efficiency in face-to-face interaction with followers (Lukaszewski & von Rueden, 2015; von Rueden et al., 2014). This tradeoff is more pronounced during situations when free-riding and coordination failure are probable or pose an existential threat to the group, such as during periods of heightened intra- or inter-group conflict."

Lines 351-357: you might comment on role of group selection models in driving hierarchy. For example as discussed in Makowsky and Smaldino (2016).

Lines 358-365: this result likely depends heavily on reproductive skew and fidelity of influence inheritance. In actual human groups, differential reproduction of influential individuals has been demonstrated for societies of all levels of hierarchy (see von Rueden and Jaeggi, 2016) so has this made humans more homogeneous in terms of influence? There may also be frequency-dependent

mechanisms maintaining differences in influence. Or mutation is stronger than you model. Or horizontal inheritance matters: the effects on influence from one's social network connections (see this demonstrated in von Rueden et al. 2019).

Lines 383-385: this though depends on precisely how the model links listening group size to time to consensus. If leader-follower interactions happened through other mechanisms, such as simultaneous influence attempts across listeners, this result may be qualitatively different, right? Worth mentioning because it's a central assumption of the model affecting the paper's main conclusions.

Lines 429-430: how was hierarchy modeled differently to make it "institutionalized", relative to the hierarchy modeled in the current study?

Lines 434-437: another key aspect missing is the collective action problem, i.e. costs from free-riding. Scalar stress might involve not just time to consensus, represented as an opportunity cost in your evolutionary model, but also free-riding or conflicts over resource division. Other models of evolution of leadership focus on these costs, e.g. Hooper et al. 2010.

Lines 450-460: per your call for more detailed study of cost of organization in groups, see work on the Tsimane comparing group size, group density, and conflict frequency to particular traits of leaders (von Rueden et al. 2014) and, most relevant to your model, to skew in political influence (Glowacki and von Rueden 2015). I don't know of other ethnographic data (particularly where hierarchies are not institutionalized) that quantitatively compares intra-group skew in political influence to metrics of scalar stress.

Alberti, G. (2014). Modeling group size and scalar stress by logistic regression from an archaeological perspective. *PLoS ONE*, 9, e91510.

Garfield, Z., von Rueden, C., and Hagen, E. (2019). The evolutionary anthropology of political leadership. *The Leadership Quarterly*, 30, 59-80.

Gavrilets, S. (2015). Collective action problem in heterogeneous groups. *Phil. Trans. Roy. Soc. B*, 370, 20150016.

Glowacki, L. and von Rueden, C. (2015). Leadership solves collective action problems in small-scale societies. *Phil. Trans. Roy. Soc. B*, 370, 20150014.

Hooper, P., Kaplan, H., and Boone, J. (2010). A theory of leadership in human cooperative groups. *Journal of Theoretical Biology*.

Olson M. (1965). *Logic of collective action: public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Makowsky M., and Smaldino, P. (2016). The evolution of power and the divergence of cooperative norms. *Journal of Economic Behavior & Organization*, 126, 75-88.

von Rueden et al. (2014). Leadership in an egalitarian society. *Human Nature*, 25, 538-566.

von Rueden, C. and van Vugt, M. (2015). Leadership in small-scale societies: some implications for theory, research, and practice. *The Leadership Quarterly*.

von Rueden, C. and Jaeggi, A. (2016). Men's social status and reproductive success in 33 non-industrial societies: effects of subsistence, marriage system, and reproductive strategy. *Proc. Natl. Acad. Sci. USA*, 113, 10824-10829.

von Rueden, C., Redhead, D., O'Gorman, R., Kaplan, H., and Gurven, M. (2019). The dynamics of cooperation and social status in a small-scale society. *Proceedings of the Royal Society B*, 286, 20191367.

Zhou, W., Sornette, D., Hill, R., and Dunbar, R. (2015). Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society B*, 272, 439-444.

Referee: 2

Comments to the Author(s)

In this paper the authors present mathematical models aimed at explaining the evolution of centralized decision-making or hierarchy in humans. This is an important topic because this is one of the most striking trends in human history yet relatively little work has been done in the way of developing formal mathematical models of the processes involved.

First an opinion formation model is developed. In this set-up individuals take turns to influence others in a group to take on their preferred value of a trait. Success in the task is measured as time taken for the group to reach consensus. Social structure is introduced by the fact that different individuals can be given different levels of influence in the model. For simplicity two classes of individuals are used in the analysis – leaders (greater influence in discussions) and followers. The initial results indicate that including hierarchy reduces the time taken for groups to make decisions. Where leaders have a strong influence then having a single leader is better as the group only gets pulled to one value.

The authors then establish an evolutionary version of the model. Pay-offs depend on how long the task takes to complete and the distribution of leaders and followers (leaders can receive a greater pay off than followers) – therefore the idea is that leaders may emerge because they enable groups to reach consensus earlier and thus receive a higher pay off. The “influence” parameter is modelled as property of individuals and is allowed to evolve. The model is explored in an island model with different groups on different patches. An additional parameter (d) is added that affects whether those with more influence do indeed receive a greater benefit.

Apologies if I have got any of that wrong but there are quite a lot of moving parts here and the manuscript is not always laid out in a way that makes it easy to follow. Overall I think the model developed here has some merit and could be published. The results seem to make intuitive sense and follow logically from the description of the model. I would like to see some steps taken to improve its clarity and also allow readers to assess its applicability to the phenomena it is interested in explaining. I want to stress that I don't think the authors have done a bad job of writing the paper it is more that the complexity of what they have done makes it challenging to take in and more could be done to break things up and help the reader.

One thing that is challenging is that terms are not always clearly defined or are introduced suddenly without being set-up sufficiently earlier in the paper. For example I am still not entirely sure what the difference the authors make here between informal hierarchy and institutional hierarchy and whether this matters for the model.

For example lines 311-316 state “The results shows that despite the wide range of possible distribution of influence, individuals evolve towards hierarchy, i.e. a minority of leaders with high influence and a majority of followers with low influence. In the meantime, the population grows to a large group size. Within a patch, informal hierarchy also evolves but the proportion of leaders and followers vary.”

It is unclear whether “informal hierarchy” within a patch is somehow different to the evolution of hierarchy described in general.

Similarly the authors make a distinction between hierarchy and inequality towards the end of the paper without really making it clear that this was an important point earlier in the paper. Calling “ d ” the level of ecological inequality is a bit confusing, and introducing ecological variation across patches as well as all the other things at the same time makes it difficult for the reader to understand the importance of any particular process.

It seems at the beginning of the paper that “scalar stress” is a key thing about the model being created yet the definition is very brief “The “iron law of oligarchy” proposes that a key element is scalar stress, which describes the fact that the cost of organisation increases with group size.”

(line 45) It would be good to give a fuller description and highlight with a specific example.

Related to this on line 377 or a result about scalar stress is introduced as a secondary finding and this was a little surprising as it seemed that the paper had been discussing this in terms of the effect of leaders in reducing time to consensus – so I am now a bit puzzled by what scalar stress actually is.

In the evolutionary model I was unclear how alpha as an individual trait is used to assign influence in the group (presumably individuals in a group are ranked according to their alpha values and then weight of influence is assigned somehow – but it wasn't clear to me from the text how this was done) – I then found the explanation earlier in the paper – however at this point a simple ascription of followers and leaders had been used so it seemed to make less sense to have that explanation there. The role of listener was also a little unclear at first.

To improve the paper I have two main recommendations; 1) include a figure and/or schematic diagrams that illustrate the model clearly and help orient the reader, 2) Break up the paper into more manageable chunks – use sub-headings. Build up the results by showing the effects of new aspects of the model (e.g. can you show the model results without the “ecological inequality” part first and then add to show the effect it has?).

There are also some modelling decisions that could be discussed further:

“The trait α carried by individuals is now an evolving trait and is transmitted vertically from parent to offspring, e.g. by social learning, as is common in both hunter-gatherer groups [28] and modern societies [29]. (lines 192-195)

The trait seems to be inherited only by vertical inheritance and is in effect no different from a genetic trait – except with a larger mutation rate (see paper by Strimling and McElreath in *Current Anthropology*). Would expect any differences if you had true cultural transmission? More could be said about the mechanism by which hierarchy spreads between populations and what selective processes are at play here (is there a balance between group and individual selection? This is implied but would be good to make more overt.) To what extent does the modelled process match what goes on in the real world?

The only benefit leaders have in this model is their ability to get individuals in a group to coordinate quickly i.e. there is no difference in outcomes or pay-offs for the particular value the group settles upon. What situations in the real-world does this apply to? What differences might we see if there were different outcomes for different values of the trait?

Related to the last point the applicability and scope of the model would be easier for the reader to understand what situations the model most closely resembles. Rather than being a general model of the evolution or importance of hierarchy it sounds like the authors are modelling something like the emergence of “Big Men” style societies where certain charismatic individuals attempt to influence members of their group. In the literature this often referred to as “achieved” rather than “ascribed” leadership because there is no office of leader that must be filled and the role is not passed on in an hereditary manner. To fit with the broad scope of PRSB it might be worth saying to what extent this model could apply to non-human animals too and provide examples.

Minor points

Lines 13-14 The independent emergence of hierarchy and its pervasive presence suggest an evolutionary origin [4].

Unclear what you mean by evolutionary origin here

Lines 19-22 “Humans have inherited traits and preferences towards hierarchy from their primate ancestors, who were organised in dominance hierarchies where individuals physically compete for rank, resources and partners [2]”

But this runs counter to the egalitarian ethos that hunter-gatherers are well-known for and even western people in lab studies have been shown to have inequity aversion

Referee: 3

Comments to the Author(s)

An impressive paper that joins an opinion formation model permitting influential leaders to emerge, thus limiting degree of scalar stress (promoting efficiency of decision-making) with an evolutionary model in which social organization is cleverly described along a continuum of skew in degree of influence, thus linking the two models. The bottom line is that the organizational advantages of leadership sufficiently limit opportunity costs and permit increasing returns to scale to permit leadership/followership to emerge even when followers are disadvantaged by lower reproduction. Leadership also allows group sizes to grow, allowing returns to scale to be realized.

General comments: this is an exceedingly complex model and anything the authors can do to help the reader fit it all together would be appreciated. For example, it might be helpful to have a single table where all the parameters are defined. The potential parameter space is very large and is only lightly sampled. (I could not see the supplementary materials however as I explained to the editor.)

I had trouble interpreting several of the key figures. On 1 & 2, what is n_{sub1} ? The population structure has not (yet at least) been explained. One assumes that n_{sub1} is the size of the first group, but how does that relate to the x axis where group size ranges from 0 to 1000? Similar issue with 2; how can groups of 30 have 50 leaders? Also 2, I presume that the linear regression coefficient depicted is the slope (should be explicit).

On 4, I wonder if it might be mislabeled. Is it the mean of alpha or its mean skewness that is indexed by the red/blue scale? Does Mean size by patch really indicate Mean group size by patch? What is meant by "Proportion of individuals in the population"? I think what we're supposed to see is that through time skewness of alpha goes up, as does mean group size, but I can't decode it.

On Figure 5C, blue, are those resources per capita, or simply in total? It seems to me that it's the per capita figures that would be critical.

General comment: since leadership has somewhat different effects here (decreasing opportunity costs as group sizes grow, thus allowing group sizes to grow which in turn allows them to enjoy increasing returns to scale) it's hard to parse at any point which of these effects dominates. Most likely they are co-constitutive and equally important, but I couldn't glean that from the ms.

Starting in 287: "On the one hand, an efficient hierarchical organisation allows a group to produce a larger amount of resources and hence reach a larger size. On the other hand, hierarchy provides a stronger advantage as group size increases because the cost of organisation increases less in hierarchical groups than in egalitarian groups." What (in the model) if anything ever puts the brakes on this process? What (in the reference societies) might do this?

"Economies of scale" are mentioned a number of times, e.g., "We make the assumption of economy of scale in which additional participants increase the benefit superlinearly [31]." I think the more common usage is to speak of increasing returns to scale, and I think that's what's actually meant (given the reference to superlinear).

There are some missed opportunities to link to other relevant work. For example, in a very recent issue of *Antiquity*, Amy Bogaard and co-authors show why agriculture where returns are limited by labor is less likely to produce wealth inequality than is agriculture limited by land. (I'm traveling and don't have access to the exact citation.) This seems (superficially at least) to be at variance with the statement "The benefit of hierarchy depends on the time constraints C_t , which translates the time to reach consensus into an opportunity cost of organisation. ... For a low level of time constraints, the benefit of hierarchy has a negligible effect on organisation and group production and thus, hierarchy does not evolve." Why this disjunction? Interesting since both of these statements ostensibly target Neolithic societies.

Are there any empirical data available that could inform us as to whether the levels of alpha used in the model for informal hierarchy are realistic? If so, please cite.

Also, quite a few uncited works come to generically similar conclusions about group size and leadership. Even though these do not (so far as I can recall) contain opinion formation models, it may be that some of these suggest that such a model is not strictly necessary? For example: Crabtree, Stefani A., R. Kyle Bocinsky, Paul L. Hooper, Susan C. Ryan, and Timothy A. Kohler 2017 How to Make a Polity (in the central Mesa Verde region). *American Antiquity* 82(1):71-95. (See also references therein.)

Also, you cite a working paper that has now been published, as follows:

Hooper, Paul, Eric Smith, Henry Wright, Timothy Kohler, and Hillard Kaplan 2018 Ecological and Social Dynamics of Territoriality and Hierarchy Formation. In *The Emergence of Pre-Modern States: New Perspectives on the Development of Complex Societies*, edited by Jeremy Sabloff and Paula L. W. Sabloff, pp. 105-132. Santa Fe Institute Press, Santa Fe.

Trivia: typos of some sort in lines 107, 444; Ref 12 incomplete.

Author's Response to Decision Letter for (RSPB-2019-2259.R0)

See Appendix A.

RSPB-2020-0693.R0

Review form: Reviewer 1

Recommendation

Accept as is

Scientific importance: Is the manuscript an original and important contribution to its field?

Excellent

General interest: Is the paper of sufficient general interest?

Excellent

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

Excellent

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

Yes

Is it clear?

Yes

Is it adequate?

Yes

Do you have any ethical concerns with this paper?

No

Comments to the Author

I thank the authors for their attention to my comments. The paper is more reader-friendly and I look forward to seeing its influence on academic debate regarding evolution of hierarchy.

Review form: Reviewer 2

Recommendation

Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?

Excellent

General interest: Is the paper of sufficient general interest?

Good

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

Good

Is the length of the paper justified?

Yes

Should the paper be seen by a specialist statistical reviewer?

No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?

N/A

Is it clear?

N/A

Is it adequate?

N/A

Do you have any ethical concerns with this paper?

No

Comments to the Author

I am satisfied with the amendments and clarifications the authors have made to the manuscript and think that the paper is now more understandable as a result.

I have only a few small points to make that I think should be addressed prior to publication:

The papers mentions the "Big Man" style of leadership in "New Guinea and Polynesia" - New Guinea is OK, but Polynesia is more often characterized by formal, institutional forms of "chiefly" leadership. I suggest using the term "Melanesia" to refer to New Guinea and the surrounding area, as this is where "Big Man" style leadership is most closely associated

On line 181 suggest rather than "economic inequality" use the term "inequality in resources" (or similar) to make things more general as readers may think of "economic" in narrow terms

lines 196-198 I appreciate that the authors have tried to address my comments about the relevance of the model to real-world processes, but I found the references to "demic diffusion" and competing by differential migration confusing in the context of this model. My comments about applicability of the model relate mainly to thinking about examples from the real-world where decision-making processes and contexts in which reaching consensus are a) important, and b) well described my the mechanism explored in this model. "Demic diffusion" is usually discussed in situations of population *expansions* where by a trait increases population growth and populations begin to spread from the initial source of innovation. I would normally associate it with a large-scale spatial trend in the spread of a trait - which does not seem to be the case being modelled here. The authors rightly note that demic diffusion has been linked to the spread of agriculture but the model is not about the spread of agriculture. So even if in some sense the model could e described as representing demic diffusion, I think this may confuse readers.

I understand that constraints of space make it difficult to include the schematic diagram in the main paper but I think this would still be useful in the supplementary materials.

Decision letter (RSPB-2020-0693.R0)

27-Apr-2020

Dear Mr Perret

I am pleased to inform you that your manuscript RSPB-2020-0693 entitled "From disorganised equality to efficient hierarchy: How group size drives the evolution of hierarchy in human societies" has been accepted for publication in Proceedings B.

The referee(s) have recommended publication, but also suggest some minor revisions to your manuscript. Therefore, I invite you to respond to the referee(s)' comments and revise your manuscript. Because the schedule for publication is very tight, it is a condition of publication that you submit the revised version of your manuscript within 7 days. If you do not think you will be able to meet this date please let us know.

To revise your manuscript, log into <https://mc.manuscriptcentral.com/prsb> and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision. You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript and upload a new version through your Author Centre.

When submitting your revised manuscript, you will be able to respond to the comments made by the referee(s) and upload a file "Response to Referees". You can use this to document any changes you make to the original manuscript. We require a copy of the manuscript with revisions made since the previous version marked as 'tracked changes' to be included in the 'response to referees' document.

Before uploading your revised files please make sure that you have:

- 1) A text file of the manuscript (doc, txt, rtf or tex), including the references, tables (including captions) and figure captions. Please remove any tracked changes from the text before submission. PDF files are not an accepted format for the "Main Document".
- 2) A separate electronic file of each figure (tiff, EPS or print-quality PDF preferred). The format should be produced directly from original creation package, or original software format. PowerPoint files are not accepted.
- 3) Electronic supplementary material: this should be contained in a separate file and where possible, all ESM should be combined into a single file. All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

- 4) A media summary: a short non-technical summary (up to 100 words) of the key findings/importance of your manuscript.

- 5) Data accessibility section and data citation

It is a condition of publication that data supporting your paper are made available either in the electronic supplementary material or through an appropriate repository.

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should be fully cited. To ensure archived data are available to readers, authors should include a 'data accessibility' section immediately after the acknowledgements section. This should list the database and accession number for all data from the article that has been made publicly available, for instance:

- DNA sequences: Genbank accessions F234391-F234402
- Phylogenetic data: TreeBASE accession number S9123
- Final DNA sequence assembly uploaded as online supplemental material
- Climate data and MaxEnt input files: Dryad doi:10.5521/dryad.12311

NB. From April 1 2013, peer reviewed articles based on research funded wholly or partly by RCUK must include, if applicable, a statement on how the underlying research materials – such as data, samples or models – can be accessed. This statement should be included in the data accessibility section.

If you wish to submit your data to Dryad (<http://datadryad.org/>) and have not already done so you can submit your data via this link

[http://datadryad.org/submit?journalID=RSPB&manu=\(Document not available\)](http://datadryad.org/submit?journalID=RSPB&manu=(Document+not+available)) which will take you to your unique entry in the Dryad repository. If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link. Please see <https://royalsociety.org/journals/ethics-policies/data-sharing-mining/> for more details.

6) For more information on our Licence to Publish, Open Access, Cover images and Media summaries, please visit <https://royalsociety.org/journals/authors/author-guidelines/>.

Once again, thank you for submitting your manuscript to Proceedings B and I look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Sincerely,
Professor Gary Carvalho
mailto:proceedingsb@royalsociety.org

Associate Editor
Board Member

Comments to Author:

Many thanks for submitting a revised version of your manuscript, and for your careful and detailed consideration of the reviewer's comments.

I agree with both reviewers that this revised version has significantly improved in terms of its understandability. Indeed, as I mentioned previously, the manuscript, and the introduction in particular, is very well written, providing interested but non-expert readers a good overview of the topic and how this study adds to our understanding of the evolution of hierarchies. At the same time, the significant changes you have made throughout the manuscript mean that also specialist readers will be able to fully appreciate your work.

While I agree with the reviewers that linking these finding to hierarchies in non-human animal societies would have been very interesting indeed, I appreciate that this would be very difficult to do while keeping within the page limit. Indeed, it would require shortening other sections, which would come at a cost in terms of its accessibility. Furthermore, even if the focus is on human societies, the discussion does a good job at placing the findings in a broader (human) context. Also, I believe that despite its focus on human societies, the manuscript will be readily accessible to those researchers with an interest in the evolution of hierarchies in other species. Thereby this manuscript will be of interest to the broader readership of Proceedings B.

Whereas reviewer 1 has no further comments, reviewer 2 has a few more, relatively minor, comments that should be easy to address. Reviewer 2 also suggests making the code available, which is something I would strongly encourage the authors to consider. Although in theory all simulations are reproducible thanks to clear and detailed descriptions, access to the code would allow the interested reader to learn more about the practical implementation.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).

I thank the authors for their attention to my comments. The paper is more reader-friendly and I look forward to seeing its influence on academic debate regarding evolution of hierarchy.

Referee: 2

Comments to the Author(s).

I am satisfied with the amendments and clarifications the authors have made to the manuscript and think that the paper is now more understandable as a result.

I have only a few small points to make that I think should be addressed prior to publication:

The papers mentions the "Big Man" style of leadership in "New Guinea and Polynesia" - New Guinea is OK, but Polynesia is more often characterized by formal, institutional forms of "chiefly" leadership. I suggest using the term "Melanesia" to refer to New Guinea and the surrounding area, as this is where "Big Man" style leadership is most closely associated

On line 181 suggest rather than "economic inequality" use the term "inequality in resources" (or similar) to make things more general as readers may think of "economic" in narrow terms

lines 196-198 I appreciate that the authors have tried to address my comments about the relevance of the model to real-world processes, but I found the references to "demic diffusion" and competing by differential migration confusing in the context of this model. My comments about applicability of the model relate mainly to thinking about examples from the real-world where decision-making processes and contexts in which reaching consensus are a) important, and b) well described my the mechanism explored in this model. "Demic diffusion" is usually discussed in situations of population *expansions* where by a trait increases population growth and populations begin to spread from the initial source of innovation. I would normally associate it with a large-scale spatial trend in the spread of a trait - which does not seem to be the case being modelled here. The authors rightly note that demic diffusion has been linked to the spread of agriculture but the model is not about the spread of agriculture. So even if in some sense the model could be described as representing demic diffusion, I think this may confuse readers.

I understand that constraints of space make it difficult to include the schematic diagram in the main paper but I think this would still be useful in the supplementary materials.

Author's Response to Decision Letter for (RSPB-2020-0693.R0)

See Appendix B.

Decision letter (RSPB-2020-0693.R1)

04-May-2020

Dear Mr Perret

I am pleased to inform you that your manuscript entitled "From disorganised equality to efficient hierarchy: How group size drives the evolution of hierarchy in human societies" has been accepted for publication in Proceedings B.

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

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

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

Your article has been estimated as being 10 pages long. Our Production Office will be able to confirm the exact length at proof stage.

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>.

Paper charges

An e-mail request for payment of any related charges will be sent out shortly. 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.

You are allowed to post any version of your manuscript on a personal website, repository or preprint server. However, the work remains under media embargo and you should not discuss it with the press until the date of publication. Please visit <https://royalsociety.org/journals/ethics-policies/media-embargo> for more information.

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,

Editor, Proceedings B

<mailto:proceedingsb@royalsociety.org>

Appendix A

Response to editor and reviewers - From disorganised equality to efficient hierarchy: How group size drives the evolution of hierarchy in human societies

We thank very much the editor and the reviewers for their detailed reviews of the previous version of the manuscript. Please find below our response to the editor and reviewers' comments, which includes the list of modifications done to the manuscript.

The reviewers' comments are shown in bold and citations to changes in the manuscript are shown in italic.

Editor

Many thanks for submitting your Proceedings B, and apologies for taking a long time to get back to you with a decision. Your patience is very much appreciated and I hope you feel it was worth the wait. I have received three excellent reviews from experts in fields directly related to the topic of your study, all three of which provide a range of insightful and constructive comments and suggestions. While they all agree that this is an impressive piece of work that provides a number of new and important insights into the evolution of hierarchies, they also all comment on the complexity of the model and express a need for some additional work to ensure your work reaches the broad readership it deserves. I fully agree with this assessment and I would therefore urge you to take this, and all other issues raised, onboard as best as you can. As a matter of fact, I found the manuscript, and the Introduction in particular, very well written. It does a very good job at setting the scene for what is to come, and to introduce the questions at hand to a non-specialist audience. This combined with the fact that you are aiming at explaining an intriguing question that is likely to appeal to many, I think there is the potential for a manuscript that will be of interest to the broad readership of Proceedings B, including those readers that would not necessarily read theoretical papers otherwise.

We thank the editor for its encouraging comments and for finding helpful and expert reviewers. We have integrated most of the modifications suggested as described below. We made an effort to clarify the manuscript. This includes adding a paragraph clearly defining terms (in particular, hierarchy and inequality), removing terms when they were not useful (e.g. informal and institutional), separating the analysis of the evolutionary model in different sections, adding sentences to explain the scope of the model, adding references and adding a paragraph at the end of the result which summarises the evolutionary dynamics in play. The detailed list of modifications done to the manuscript can be found below.

Referee: 1

I think the model will be of broad interest and favor publication. The model makes a number of predictions that have potential to drive much future research. In fact, the discussion would benefit from more attention to specific predictions the model makes (e.g. how # of leaders interacts with leader-follower influence gap, or how time constraints affect hierarchy evolution) and how these predictions comport with existing ethnographic/experimental data. I also think the authors should in places amend or clarify how they frame the model and how they discuss its implications (see comments below). In particular, the way the authors model how leaders and followers influence each other in groups (the opinion formation model) is central to their findings that hierarchy reduces scalar stress and that hierarchy can evolve on that basis. Since so much rests on this modeling decision, I think more caveats are needed or attention to how such leader-follower interaction may or may not mirror actual group decision-making. Also, the focus is on humans, but why not consider how the model may inform evolution of hierarchy in other species, particularly for the Proc B audience?

We have modified the manuscript to answer these comments (please see below). However, there are two main comments that we can not integrate in this manuscript. First, we agree with the reviewer on the importance of discussing more extensively how predictions of the model compare with empirical data. Nevertheless, this is not possible considering the length requirements of Proceedings of the Royal Society B and the complexity of this model which already requires a detailed description. This comparison will be done in future work where we would like to test the predictions of the model on empirical data. Similarly, we agree that a more extensive discussion of the assumptions and caveats of the opinion formation model would be useful. For similar reasons than before, this can not be done in this manuscript but the opinion formation model has been more extensively analysed and discussed in a paper in preparation for publication.

- **Title: “disorganized equality or efficient despotism” makes it**

sound like you're modeling hierarchy as a binary, and "despotism" implies coercion (which you're also not modeling per se). Maybe change wording here.

We modified the title to:

"From disorganised equality to efficient hierarchy: How group size drives the evolution of hierarchy in human societies"

We added the terms "from" and "to" to integrate the comment on social organisation being modelled as a continuous range rather than in a binary form. We changed "despotism" to "hierarchy" to avoid implying coercion.

- **Line 10: Boehm's focus is on the absence of dominance-based hierarchy, and he de-emphasizes or omits the prestige-based hierarchy that often does influence informal political decision-making in these societies. See Garfield et al. (2019).**

We added the suggested reference to provide a more complete view of the literature on this topic.

- **Lines 15-18: Framing leaders' preferential access to resources or mating partners as exploitation is misleading, to the extent this results at least in part from voluntary exchange. Furthermore, leaders likely benefited reproductively throughout human evolution, even where there were minimal material resources to contest and an egalitarian ethos limited the extent of hierarchy. See meta-analysis by von Rueden and Jaeggi (2016).**

We removed the term of exploitation from the sentence and replaced it by "if the position of leader is more beneficial" 117:

"Why would any individual rationally accept a position of being a follower if the position of leader is more beneficial?"

- **Lines 22-26: "equality" is misleading. Differences in informal political influence according to prestige or gender or age were likely commonplace.**

We modified the sentence cited and replaced "maintained equality" by "avoided coercive leaders" 125:

"In pre-Neolithic tribes, coalitions of followers imposed strong dominance on uprising leaders and successfully avoided coercive leaders for hundreds of thousand of years [1]."

- **Furthermore, sustaining relative egalitarianism may counter-intuitively depend in part on motivations to acquire status via cooperation, not just anti-dominance behavior by followers. You only describe the latter. When status depends on demonstrating value to others (such as through effective leadership), this can result in transfer of knowledge, resources, or reputation from higher status to lower status individuals, limiting growth in status skew**

over time. See von Rueden et al. (2019) for an empirical demonstration of this using longitudinal ethnographic data.

We agree with the reviewer that egalitarianism can also be explained by other causes than coercion. We cite only coercion here because the aim of this paragraph is to explain why coercion is not enough to explain the emergence of permanent hierarchy at the Neolithic transition, rather than presenting exhaustively the causes behind the emergence of egalitarian organisation. This is shown by the preceding sentence:

“One could argue that followers do not have a choice because leaders impose their dominance using coercive means.”

- **Lines 38-40: these claims need citation. There is a behavioral economics and a social psychology literature that addresses effects of leaders on things like time to consensus or steering their group towards a cooperative equilibrium, via “first mover” effects or punishment/reward capability.**

We added references which describe the numerous roles of leaders group organisation l37:

“As can be seen in small scale societies [2] or in a modern share-holder meeting [3], leaders reduce the cost of organisation by assigning roles to individuals, settling arguments between decision makers, and helping to decide the future course of action.”

We also added one reference on the effect of leaders on consensus time l68: *“The iron law of oligarchy [4] and behavioural experiments [5] suggest that the benefit of hierarchy on group coordination lies in its effect on the time a group spends to reach consensus and take a collective decision (consensus time in short).”*

- **Lines 41-44: there can also be a conflict over who adopts leader or follower roles, or lack of motivation to adopt leader role at all given its costs- so introducing inter-individual differences into expected costs and benefits of leading may be necessary. See Gavrillets (2015), which takes inspiration from Olson (1965). The traits you later mention, including personality and body size, can affect these costs and benefits.**

Although this would be an interesting development, our research questions in this paper are not focused on who in particular takes the leader roles and the effect of inter-individual differences on their motivations. Consequently, while we acknowledge that there can be a lack of motivation for individuals to adopt the leader role, we restrain our model to cases where leaders receive a higher share of resources that offsets the costs of leading.

- **Lines 45-56: what is meant by scalar stress? Increasing inability to reach consensus/coordinate? Increasing dyadic conflict/free-riding on collective action? Your model focuses on the former,**

but define scalar stress here because that helps reader anticipate what your mechanistic model will explain. You might cite here Alberti (2014), who fits a simple logistic model to data from Hutterites on group size and fissioning, which he argues has broad cross-cultural validity as a metric of “scalar stress”. Though it doesn’t analytically model the mechanisms that may affect the relationship between group size and fissioning.

We have added a more precise definition of scalar stress which explains that we referring to coordination rather than cooperation/free-riding, and have cited the paper proposed by the reviewer l45:

“The “iron law of oligarchy” proposes that a key element is scalar stress, which describes the fact that the difficulty for a group to coordinate increases with group size [4]. This relation appears in (i) psychology experiments of collective decision-making, in which larger groups reach a lesser degree of agreement [6] or take worse decisions [7], and (ii) indirectly in anthropological data showing a strong correlation between group size and probability of group fission [8] or group size and the number of political units [9]. ”

We do not give more details at this point of the manuscript because the paragraph aims to highlight the lack of formal model of scalar stress. We define formally scalar stress later in the introduction l93:

“We investigate how social organisation affects scalar stress, where scalar stress is defined as the relationship between the time spent to reach consensus and group size. ”

- **Lines 58-63: the discussion of inter-individual differences affecting emergence of informal hierarchy feels like a non-sequitur here. Inter-individual differences inform leader emergence where hierarchy is institutionalized as well. Indeed, your citations here are of studies based on leader emergence in the context of institutionalized hierarchy. [...]**

The goal of this paragraph is to explain one gap in literature, which is the absence of a model investigating the emergence of hierarchy from the evolution of individual characteristics. We have modified this paragraph to make it more clear l55:

“Previous work has shown that scalar stress can drive the evolution of institutionalised hierarchy [10], where a leader is appointed by a centralised process. However, rather than being ascribed, hierarchy is likely to initially emerge from the evolution of intrinsic physical and psychological traits of individuals, e.g. height [11], talkativeness and charisma [12]. Emblematic examples of such informal hierarchy are the “big man” societies observed in Polynesia and New Guinea, in which leaders are defined by their persuasion skills rather than by an ascribed position [13]. But explaining the evolution of “informal” hierarchy without supporting institutions poses an important challenge: can the group benefit of hierarchy overcome the selec-

tion pressure pushing everyone to be a leader? The lack of a mechanistic model describing the effect of hierarchy on collective decision-making has limited investigation of scalar stress as a possible solution.”

- **Lines 58-63:** [...] Furthermore, inter-individual differences themselves don't explain emergence of informal hierarchy. You would need to describe how these differences affect acquisition of reputations for an ability to benefit others (prestige) or harm others (dominance), leading to deference to the prestigious or dominant (i.e. hierarchy)- or, as in your model, becoming a speaker and being influenced during speaking events.

We do not integrate this reputation mechanism in the manuscript because we are focusing on ultimate explanations for the emergence of hierarchy. We put in evidence selection pressures (an ultimate explanation) affecting the evolution of influence, with influence being described as the capacity to spread your opinion. Proximately, this influence can result from particular traits or behaviours. As suggested by the reviewer, influence can describe individual investment in gaining reputation, which subsequently results in better spread of its opinions. However, influence could also describe the verbal ability of individuals, which also results in a better spread of their opinions. The ultimate approach explored in this manuscript has for benefit to apply to any of these mechanisms. However, it is important that future work explore proximate mechanisms by which influence and hierarchy emerge, e.g. the effect of reputation dynamics.

- **Lines 75-78:** again, be careful implying an opposition between egalitarianism and leader/follower behavior, without more qualification. Also, before you present these questions, define what you mean by inequality, informal hierarchy, and leadership. You use these terms in discussing the specification and results of your models, but it's not always clear how they differ. Hierarchy to many readers will imply differences in access to contested resources, but hierarchy has also been defined in terms of differential influence (as revealed in deference signaling between group members). In your models it appears you use the latter definition, but I don't think this was clear until later in the paper (e.g. Line 284). In your opinion formation model, the distribution of leadership is tantamount to hierarchy, because you require that leaders (i.e. speakers) always have greater influence. Furthermore, greater influence results in inequality: group decisions are more favorable to leaders (i.e. closer to their preference) than the average group member. However, in the evolutionary model, inequality is distribution of production, which is distinct from, though a function of, group members' influence and interests.

We have reworded the terms used. We now call groups without hierarchy (a skewness of influence close to 0) *acephalous* rather than egalitarian. We

describe the inequality in term of access to resources as *economic inequality*. We have modified a paragraph in the introduction to introduce these definitions 174:

“We describe group social organisation as a distribution of individuals’ influence, i.e. their capacity to modify another individual’s opinion toward their own. The scale from acephalous to highly hierarchical groups is represented by an equal to strongly-positively-skewed distribution of influence. We use the term leaders and followers to describe individuals with high and low influence, respectively. This definition of hierarchy does not include the degree of economic inequality, which we define as the skewness of the distribution of resources. We allow the correlation between hierarchy and degree of economic inequality to emerge from the model, as influential individuals can bias the distribution of resources to their advantage. The emergence of hierarchy is represented by the evolution of individual behaviours towards a minority of leaders and a majority of followers.”

- **Also, in the evolutionary model, hierarchy (as you define it) can coexist with zero inequality (when $d=0$). In other words, leaders might be instrumental in coordinating a final decision that is close to their interests even where the final decision doesn’t differentially benefit them. Right? If so, does it make sense to have a group decision that is close to one’s interests yet doesn’t differentially benefit you? This is possible if leaders experience favoritism in subsequent exchange on account of their effective leadership (i.e. their gain in prestige). This more commonly describes the benefits to leadership in more egalitarian settings.**

In the case with zero inequality ($d = 0$), leaders receive the same resources than the members of their group but they get a benefit compared to members of other groups without hierarchy. As underlined by the reviewer, the case of perfect equality within the group might be rare in the real world, which is why most of the results presented consider inequality ($d = 1$). However, the case with zero inequality is useful to understand the effect of inequality on the dynamics of the model. We agree with the reviewer that future models could look at more complex proximate mechanisms by which leaders might gain a higher amount of resources, such as exchange and prestige.

- **Line 118: define k**

We added a sentence to define k 1126:

“The exponent k defines how much the difference in influence is translated into a difference in the probability to talk.”

- **Line 124: Dunbar has more recent work that may be even more relevant to this claim, e.g. Zhou et al. 2015.**

We thank the reviewer for pointing out this publication. However, we removed the reference to Dunbar. This is because we removed the justifica-

tion that the number of listeners is limited because of cognitive constraints limiting the number of relationships that an individual can maintain. This type of cognitive constraint explains the limit on the number of long-term social relationships, while the model looks at short term interactions. This is also why the network is complete. The sentence is now l130:

“The number of listeners is limited because of time constraints.”

- **Lines 129-131: If the listener is more influential, equation 2 can mean the listener’s preference moves even further away from the speaker’s preference. For example, $a(u)=.5$, $a(v)=1$, $x(u)=.5$, and $x(v)=1$. Then $x'(v)=1.25$. Implementing a rule that speakers are always more influential than listeners by a minimum value (as you do) circumvents this but (a) seems artificial (listeners who are more influential are arbitrarily made less influential than speakers) and (b) cooks hierarchy into the results even more. Influence is already affecting the probability one’s own opinion will ever change (based on equation 1).**

We agree that this minimum value implemented in the model is artificial. Alternative update functions can avoid this problem, but give rise to other biases. However, we have investigated opinion formation models with other update functions and have obtained similar qualitative results. These results can not be presented here because of the length requirement but they will be presented in a future paper in preparation.

- **Line 138: Why define consensus as below a threshold standard deviation of preferences, rather than complete preference agreement? This suggests you are building compromise into your model, where individuals are willing to follow the average view so long as it’s below that threshold. Why should compromise be more likely when opinions vary little as opposed to opinions varying a lot? I’d mention this when describing the model in the introduction.**

We use a threshold standard deviation of preferences to define the consensus rather than complete preference agreement because complete preference agreement is impossible for continuous opinions. The opinions can take an infinite number of values between 0 and 1, and so the probability that they all have the same value tends to 0. Thus, we assume here that very close opinions are similar opinions (or that individuals can not see the difference between very close opinions).

- **Lines 150-151: very interesting results- particularly that single leaders (as opposed to multiple hierarchy as you call it) are only more efficient re reducing scalar stress when leader-follower influence gap is medium to high. This seems like a good opportunity to describe the importance of analytical models for generating new predictions- how does your result match experimental**

or observational data comparing number of leaders, leader influence, and group outcomes? In the discussion, you don't address fit of your specific results (rather than their broad implications) to existing data, nor call for tests of specific predictions arising from your model. I would think though that group size and leader-follower influence gap are rarely independent. Managerial mutualism models of hierarchy are built around idea that increasing group size makes group members more willing to offer greater deference or pay a higher fee to managers, to surmount increasing coordination or collective action problems.

We agree that it is important to compare the results of the model to empirical data. However, there is a limited number of studies that quantitatively test these relations and that would allow a precise test of the results. Furthermore, the length restrictions do not allow us to go beyond broad implications. This would require a more extensive study that will be done in future work.

We call for tests of predictions of the model in the discussion 1478:

“Yet, the generality of a scenario where organisation drives the evolution of hierarchy needs to be better estimated with further work exploring the quantitative relation between individual behaviours, group size and cost of organisation either in laboratory experiments or in real world human groups.”

- **Lines 171-174:** I would like to see a direct comparison of effect of group size and effect of sampling size (number of listeners) on time to consensus. The effect of sampling size may be related to the effect of group density in driving hierarchy. For example, a study of Amazonian villages suggests that centralization of leadership (i.e. how skewed influence is across individuals) may be more closely linked to residential density than to group size (Glowacki and von Rueden, 2015). In general, I would think group size only matters insofar as it indicates how often people are interacting and in what numbers and how often differences of opinion arise, something density may capture more directly.

Figure 3 presents the interactions between these two factors (group size and number of listeners). We agree that group size matters only if it describes the number of individuals interacting, and that density might be a better descriptor. We thank the reviewer for pointing out this difference, which will be particularly important when comparing the model to empirical data.

- **Lines 171-174:** Perhaps counter-intuitively, your model indicates that there is less scalar stress (lower regression coefficient) the greater the number of listeners per time point (irrespective of number of leaders). And that the number of leaders matters less

for scalar stress the greater the number of listeners. Yet intuitively more listeners would suggest more scalar stress, and that leadership should matter more in limiting scalar stress as number of listeners increases. This is a particularity of the model, in that number of listeners is not a metric of number of people trying to reconcile each others' preferences simultaneously. Would be good to emphasize/reiterate this.

As stressed by the reviewer, the number of listeners describes the number of individuals listening in a single event of discussion and is different from the total number of individual undergoing consensus decision making or total number of individuals interacting. To make it clear, we modified the following sentence to emphasise this difference l130:

"The speaker talks with N_l listeners, with listeners being a subset of the total group."

- **Figure 3 caption:** I'd stipulate that its number of leaders and number of listeners per discussion event. A casual reader might think the first box of the figure includes analysis of up to 50 leaders in a group size of only 10.

Figure 3 (now supplementary Figure 2) has been modified to be more explicit. For instance, the first box of the figure is now titled " $N_l = 10$ " rather than "10".

- **Lines 273-275:** how is the final decision (x^*j) generated? Based on the opinion formation model? If so then realized influence is in a sense how one's initial influence updates, right? Based on how well that initial influence determined the proximity of one's preference to the final decision?

The final decision is generated in the evolutionary model from the opinion formation model as defined earlier l145:

"The final decision reached, x^ , is the mean of the opinions at consensus across individuals."*

Realised influence represents how much the individual has effectively employed its influence to bring the opinion of others close to its preferences (here represented by initial opinion).

- **Lines 311-314:** Figure 4 suggests that as simulation progresses across generations, the range of influence (and possibly skew of influence) becomes much smaller, contrary to what is claimed and what is shown in Figure 5. This just may be because there are so few individuals with influence above 0.5 after 4000 generations that they aren't visible. Maybe mention this.

The reviewer highlighted that the Figure 4 can be misleading because the small proportion of leaders is almost not visible. To solve this issue, we have added a sentence in the caption of the Figure 4 (now Figure 2):

“Note that there is a small proportion of individuals with high influence at equilibrium. This proportion is low and hard to discern but it is revealed by a stripe of orange and white colour at the bottom.”

- **Lines 340-342:** the relationship between time constraints and evolution of hierarchy is another prediction of the model that could be given more time in the discussion. There is a big literature, for example, on how existential threats like war (which you suggest carries greater time constraints) affect choice of leader. Under such conditions, leaders are often granted more influence and dominance in leaders may produce less blowback. So human leader choice may have evolved as a facultative mechanism, if groups consistently experienced activities with different time constraints and thus variable effects of influence skew on fitness. See ethnographic and experimental references in von Rueden and van Vugt (2015): “Crow and Omaha war captains exercised authority during warfare that was denied them back in camp (Lowie, 1948), Yanomamo headmen from Venezuela were granted authority during raids that they lacked during other times (Chagnon, 1983), the Nambiquara of Brazil endorsed more authoritarian leadership when under threat of attack by neighboring groups (Price, 1981), and Enga leaders’ influence in highland New Guinea was heightened during periods of conflict while during peace leadership was challenged and frequently changed hands (Meggitt, 1967).” “preferences for larger or more masculine leaders may be greater when group members experience more conflict (van Vugt de Cremer, 1999) or are under threat from out-groups (Blaker Van Vugt, 2014; Halevy, Chou, Cohen, Livingston, 2012; Laustsen Petersen, 2015; Little et al., 2007; Spisak et al., 2011; Tigue et al., 2012). In conjunction with our review of leadership in SSSs, these findings suggest that humans evolved a psychology that trades-off the risk of exploitation by physically dominant leaders with their greater coordination and conflict resolution efficiency in face-to-face interaction with followers (Lukaszewski von Rueden, 2015; von Rueden et al., 2014). This tradeoff is more pronounced during situations when free-riding and coordination failure are probable or pose an existential threat to the group, such as during periods of heightened intra- or inter-group conflict.”

We thank the reviewer for pointing out these references and their empirical evidence. We have added the following sentences in the discussion 1466:

“Our findings predict that the level of hierarchy i.e. skewness of the distribution of influence, should increase both with the time constraints on the tasks tackled by the group and with group size. First, there is extensive evidence that human groups tackling tasks with high time constraints such

as warfare often switch to a strong hierarchical organisation [14].”

We share the intuition of the reviewer that human leader choice may have evolved as a facultative mechanism. Again, we can not discuss this explicitly because of length requirements but it would be a very interesting extension.

- **Lines 351-357: you might comment on role of group selection models in driving hierarchy. For example as discussed in Makowsky and Smaldino (2016).**

We added a paragraph at the end of the results which describes the evolutionary dynamics of the model l396:

“To summarise, social organisation is the equilibrium between two opposing forces, competition within groups where inequality pushes individuals to evolve high influence, and competition between individuals of different groups where efficient group organisation pushes most individuals to evolve low influence. Looking closer, hierarchy provides one direct and one indirect benefit [15] to followers compared to individuals in acephalous groups. First, hierarchy provides a direct benefit to followers because it increases the amount of surplus resources produced and thus, it increases the fitness of followers. Second, hierarchy provides an indirect benefit to followers because it increases the group size and hence the amount of resources produced in the following generation. This increases the fitness of followers’ offspring. The contribution of each benefit is hard to distinguish but their role can be examined by investigating the effect of high migration rate, which suppresses population structure and indirect benefits to offspring on the same patch. Supplementary Figure 5 shows that, considering moderate time constraints, a high migration rate leads to the disappearance of hierarchy at equilibrium. This highlights the importance of the indirect benefit to offspring that remain on the patch in sustaining hierarchy. On the other hand, supplementary Figure 6 shows that hierarchy evolves for any migration rate if the the time constraints are high. In this case, the direct benefit is high enough to overcome the cost of economic inequality. In conclusion, hierarchy can evolve when time constraints are high through the immediate direct benefit of producing extra resources, but the indirect benefit resulting from the feedback loop between hierarchy, group size and scalar stress allows hierarchy to evolve over a much wider range of conditions.”

We use the term competition between individuals of different groups rather than group selection as this term can carry different meanings among different communities. For instance, some researchers would consider that there is no group selection *stricto sensu* in this model, because groups do not replicate.

We thank the reviewer for bringing our attention to the paper of *Makowsky et al* [16]. However, we prefer to not discuss this paper here as this paper

relies on a strong assumption, which lacks empirical support: There is a correlation between individual behaviour (co-operator or defector) and opinion on the group actions (peace or war).

- **Lines 358-365: this result likely depends heavily on reproductive skew and fidelity of influence inheritance. In actual human groups, differential reproduction of influential individuals has been demonstrated for societies of all levels of hierarchy (see von Rueden and Jaeggi, 2016) so has this made humans more homogeneous in terms of influence? There may also be frequency-dependent mechanisms maintaining differences in influence. Or mutation is stronger than you model. Or horizontal inheritance matters: the effects on influence from one's social network connections (see this demonstrated in von Rueden et al. 2019).**

We agree that there is a discrepancy between theory predicting homogeneous groups in term of influence and observations of skewed distributions of influence. This discrepancy is one of the main motivations of this paper. As the reviewer suggests, this model proposes and investigates one frequency-dependence mechanism which could explain this heterogeneity in influence. We agree that other mechanisms could explain the distribution of influence observed in human societies and it is important that further works explore other mechanisms as done in the work suggested by the reviewer.

It is the case that the results depend on the fidelity of the inheritance of the trait. However, lower fidelity or higher mutation would increase noise but not remove qualitative results (no fidelity would remove any results). However, this comment applies to any evolutionary model. Finally, if the model would consider horizontal inheritance with pay-off biased learning, the qualitative result of the model would be the same [17].

- **Lines 383-385: this though depends on precisely how the model links listening group size to time to consensus. If leader-follower interactions happened through other mechanisms, such as simultaneous influence attempts across listeners, this result may be qualitatively different, right? Worth mentioning because it's a central assumption of the model affecting the paper's main conclusions.**

We agree with the reviewer that the mechanisms underlying leader-follower interactions could modify the results of the model. The opinion formation model should be expanded in further work to explore these alternatives. However, the sentence cited claims a correlation between the degree of scalar stress and the degree of hierarchy, rather than a correlation between number of listeners and degree of hierarchy. This prediction results from the evolutionary model and will still hold if scalar stress emerges via another mechanism, as observed in empirical evidence. To be clearer, we

modified the sentence l382:

“On the other hand, a low scalar stress, i.e. here represented by a high number of listeners N_1 , leads to the disappearance of hierarchy. This result shows that the benefit of reducing scalar stress is a crucial factor in the evolution of hierarchy.”

- **Lines 429-430: how was hierarchy modeled differently to make it “institutionalized”, relative to the hierarchy modeled in the current study?**

The previous model with institutionalised hierarchy models hierarchy as the result of group decision. A group has hierarchy if more than half of the individuals in the group have preferences for hierarchy. In the present manuscript, hierarchy is dependant of individuals’ traits, rather than the result of a group decision. A group has hierarchy if a minority of individuals have leader traits and a majority have follower traits. It is important because this results in different dynamics. We modified the following sentences to make the difference more explicit, first in the introduction l55:

“Previous work has shown that scalar stress can drive the evolution of institutionalised hierarchy [10], where a leader is appointed by a centralised process. However, rather than being ascribed, hierarchy is likely to initially emerge from the evolution of intrinsic physical and psychological traits of individuals, e.g. height [11], talkativeness and charisma [12]. Emblematic examples of such informal hierarchy are the “big man” societies observed in Polynesia and New Guinea, in which leaders are defined by their persuasion skills rather than by an ascribed position [13]. But explaining the evolution of “informal” hierarchy without supporting institutions poses an important challenge: can the group benefit of hierarchy overcome the selection pressure pushing everyone to be a leader? ”

And in the discussion l450:

“Our model of informal hierarchy extends previous work on institutionalised hierarchy [10] by showing that this voluntary theory holds even in societies where political institutions are absent, and thus where inequality creates selection pressures toward leader behaviours.”

- **Lines 434-437: another key aspect missing is the collective action problem, i.e. costs from free-riding. Scalar stress might involve not just time to consensus, represented as an opportunity cost in your evolutionary model, but also free-riding or conflicts over resource division. Other models of evolution of leadership focus on these costs, e.g. Hooper et al. 2010.**

We agree with the reviewer that the role of leaders in solving collective action problems could also explain the emergence of hierarchy. We cite the paper proposed by the reviewer in the discussion l459:

“More broadly, our model is in line with theoretical work that proposes that hierarchy emerged because leaders fulfil an important role for the group e.g. leaders promote cooperation by monitoring the group and punishing the

defectors [18]. Importantly, the explanation explored here is not mutually exclusive with previous explanations. Rather, it can complete them. Following the previous example [18], policing in large-scale societies requires efficient decision-making to create the large number of rules [19] and to manage specialised policing forces.”

- **Lines 450-460:** per your call for more detailed study of cost of organization in groups, see work on the Tsimane comparing group size, group density, and conflict frequency to particular traits of leaders (von Rueden et al. 2014) and, most relevant to your model, to skew in political influence (Glowacki and von Rueden 2015). I don’t know of other ethnographic data (particularly where hierarchies are not institutionalized) that quantitatively compares intra-group skew in political influence to metrics of scalar stress.

We thank the reviewer for pointing out these references which would be useful to test the results of our model. We did not integrate them in the manuscript because they look at conflict resolution rather than coordination. It is likely that these data still apply to our model because conflict resolution can be seen as a form of coordination. However, we believe that the space available in this manuscript would not allow us to give this discussion the level of details it requires, and would prefer to keep this discussion for future work.

Alberti, G. (2014). Modeling group size and scalar stress by logistic regression from an archaeological perspective. *PLoS ONE*, 9, e91510.

Garfield, Z., von Rueden, C., and Hagen, E. (2019). The evolutionary anthropology of political leadership. *The Leadership Quarterly*, 30, 59-80.

Gavrilets, S. (2015). Collective action problem in heterogeneous groups. *Phil. Trans. Roy. Soc. B.*, 370, 20150016.

Glowacki, L. and von Rueden, C. (2015). Leadership solves collective action problems in small-scale societies. *Phil. Trans. Roy. Soc. B*, 370, 20150014.

Hooper, P., Kaplan, H., and Boone, J. (2010). A theory of leadership in human cooperative groups. *Journal of Theoretical Biology*.

Olson M. (1965). *Logic of collective action: public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Makowsky M., and Smaldino, P. (2016). The evolution of power and the divergence of cooperative norms. *Journal of Economic Behavior Organization*, 126, 75–88.

von Rueden et al. (2014). Leadership in an egalitarian society. *Human Nature*, 25, 538-566.

von Rueden, C. and van Vugt, M. (2015). Leadership in small-scale societies: some implications for theory, research, and practice. *The Leadership Quarterly*.

von Rueden, C. and Jaeggi, A. (2016). Men’s social status and reproductive success in 33 non-industrial societies: effects of subsistence, marriage system, and reproductive strategy. *Proc. Natl. Acad. Sci. USA*, 113,

10824-10829.

von Rueden, C., Redhead, D., O’Gorman, R., Kaplan, H., and Gurven, M. (2019). *The dynamics of cooperation and social status in a small-scale society*. *Proceedings of the Royal Society B*, 286, 20191367.

Zhou, W., Sornette, D., Hill, R., and Dunbar, R. (2015). *Discrete hierarchical organization of social group sizes*. *Proceedings of the Royal Society B*, 272, 439-444.

Referee: 2

Apologies if I have got any of that wrong but there are quite a lot of moving parts here and the manuscript is not always laid out in a way that makes it easy to follow. Overall I think the model developed here has some merit and could be published. The results seem to make intuitive sense and follow logically from the description of the model. I would like to see some steps taken to improve its clarity and also allow readers to assess its applicability to the phenomena it is interested in explaining. I want to stress that I don’t think the authors have done a bad job of writing the paper it is more that the complexity of what they have done makes it challenging to take in and more could be done to break things up and help the reader.

We have modified the manuscript to address these comments as far as possible, and have followed the suggestions of the reviewer to improve clarity (see below).

- One thing that is challenging is that terms are not always clearly defined or are introduced suddenly without being set-up sufficiently earlier in the paper. For example I am still not entirely sure what the difference the authors make here between informal hierarchy and institutional hierarchy and whether this matters for the model. For example lines 311-316 state “The results shows that despite the wide range of possible distribution of influence, individuals evolve towards hierarchy, i.e. a minority of leaders with high influence and a majority of followers with low influence. In the meantime, the population grows to a large group size. Within a patch, informal hierarchy also evolves but the proportion of leaders and followers vary.” It is unclear whether “informal hierarchy” within a patch is somehow different to the evolution of hierarchy described in general.

We removed the term “informal” in most of cases because (i) the manuscript describes only the emergence of informal hierarchy and (ii) hierarchy as studied here is formally defined in the manuscript. We kept the term “informal” when it is used to highlight differences between this model and previous models of institutionalised hierarchy (one time in the introduc-

tion and one time in the discussion). We also modified these sentences to make the difference more explicit, first in the introduction 155:

“Previous work has shown that scalar stress can drive the evolution of institutionalised hierarchy [10], where a leader is appointed by a centralised process. However, rather than being ascribed, hierarchy is likely to initially emerge from the evolution of intrinsic physical and psychological traits of individuals, e.g. height [11], talkativeness and charisma [12]. Emblematic examples of such informal hierarchy are the “big man” societies observed in Polynesia and New Guinea, in which leaders are defined by their persuasion skills rather than by an ascribed position [13]. But explaining the evolution of “informal” hierarchy without supporting institutions poses an important challenge: can the group benefit of hierarchy overcome the selection pressure pushing everyone to be a leader?”

And in the discussion 1450:

“Our model of informal hierarchy extends previous work on institutionalised hierarchy [10] by showing that this voluntary theory holds even in societies where political institutions are absent, and thus where inequality creates selection pressures toward leader behaviours.”

- **Similarly the authors make a distinction between hierarchy and inequality towards the end of the paper without really making it clear that this was an important point earlier in the paper. Calling “d” the level of ecological inequality is a bit confusing, and introducing ecological variation across patches as well as all the other things at the same time makes it difficult for the reader to understand the importance of any particular process.**

We have reworded terms to avoid the confusion stressed by the reviewer. We now call groups with a skewness of influence close to 0 acephalous rather than egalitarian. We describe the inequality in term of access to resources as economic inequality. We modified a paragraph in the introduction to introduce these definitions 174:

“We describe group social organisation as a distribution of individuals’ influence, i.e. their capacity to modify another individual’s opinion toward their own. The scale from acephalous to highly hierarchical groups is represented by an equal to strongly-positively-skewed distribution of influence. We use the term leaders and followers to describe individuals with high and low influence, respectively. This definition of hierarchy does not include the degree of economic inequality, which we define as the skewness of the distribution of resources. We allow the correlation between hierarchy and degree of economic inequality to emerge from the model, as influential individuals can bias the distribution of resources to their advantage. The emergence of hierarchy is represented by the evolution of individual behaviours towards a minority of leaders and a majority of followers.”

The level of ecological inequality is not different across patches as stated in 1275: *“It is assumed for simplicity that d is the same for all patches, and*

is determined for example by the state of technology, such as food storage and military technologies. Nevertheless, different patches can have more or less despotic distributions of resources due to different distributions of α_r values.”

- **It seems at the beginning of the paper that “scalar stress” is a key thing about the model being created yet the definition is very brief “The “iron law of oligarchy” proposes that a key element is scalar stress, which describes the fact that the cost of organisation increases with group size.” (line 45) It would be good to give a fuller description and highlight with a specific example. Related to this on line 377 or a result about scalar stress is introduced as a secondary finding and this was a little surprising as it seemed that the paper had been discussing this in terms of the effect of leaders in reducing time to consensus – so I am now a bit puzzled by what scalar stress actually is.**

We have added a more precise definition of scalar stress which explains that we referring to coordination rather than cooperation/free-riding, and have cited the paper proposed by the reviewer 145:

“The “iron law of oligarchy” proposes that a key element is scalar stress, which describes the fact that the difficulty for a group to coordinate increases with group size [4]. This relation appears in (i) psychology experiments of collective decision-making, in which larger groups reach a lesser degree of agreement [6] or take worse decisions [7], and (ii) indirectly in anthropological data showing a strong correlation between group size and probability of group fission [8] or group size and the number of political units [9]. ”

We do not give more details at this point of the manuscript because the paragraph aims to highlight the lack of formal model of scalar stress. We define formally scalar stress later in the introduction 193:

“We investigate how social organisation affects scalar stress, where scalar stress is defined as the relationship between the time spent to reach consensus and group size. ”

The result on scalar stress is given in the latter part of the discussion to make the results simpler to understand. Indeed, scalar stress involves a feedback loop between group size and hierarchy, which is easier to understand if the reader has a grasp of the main dynamics of the model.

- **In the evolutionary model I was unclear how alpha as an individual trait is used to assign influence in the group (presumably individuals in a group are ranked according to their alpha values and then weight of influence is assigned somehow – but it wasn’t clear to me from the text how this was done) – I then found the**

explanation earlier in the paper – however at this point a simple ascription of followers and leaders had been used so it seemed to make less sense to have that explanation there. The role of listener was also a little unclear at first.

We modified the following sentence and moved it to the beginning of the definition of the opinion formation model l118:

“To facilitate the analysis of the opinion formation model, we divide individuals into two profiles: leader L, and follower F. Each profile has a fixed value of influence α such that $\alpha_L > \alpha_F$.”

We added to the definition of the evolutionary model l187:

“Unlike the model above, the possible values of influence α are now not restrained to two profiles but can take any value in the range $[0,1]$. We use the terms leader and follower to designate individuals with high and low influence, respectively.”

- **To improve the paper I have two main recommendations; 1) include a figure and/or schematic diagrams that illustrate the model clearly and help orient the reader, 2) Break up the paper into more manageable chunks – use sub-headings. Build up the results by showing the effects of new aspects of the model (e.g. can you show the model results without the “ecological inequality” part first and then add to show the effect it has?).**

We thank the reviewers for the suggestions. We created three subsections in the analysis of the evolutionary model. We have done several modifications to make the terms and definition clearer. However, the length requirements of Proceedings of Royal Society B do not allow us to add a figure.

- **There are also some modelling decisions that could be discussed further: “The trait α carried by individuals is now an evolving trait and is transmitted vertically from parent to offspring, e.g. by social learning, as is common in both hunter-gatherer groups [28] and modern societies [29]. (lines 192-195) The trait seems to be inherited only by vertical inheritance and is in effect no different from a genetic trait – except with a larger mutation rate (see paper by Strimling and McElreath in Current Anthropology). Would expect any differences if you had true cultural transmission?**

We assume that the reviewer means horizontal transmission by true cultural transmission. Horizontal transmission with pay-off biased learning [17] would carry similar qualitative results. This would also depend on how much social learning is done between individuals within a same group or with individuals outside of the group, which is analogous to the migration rate in our model.

- **More could be said about the mechanism by which hierarchy spreads between populations and what selective processes are at play here (is there a balance between group and individual selection? This is implied but would be good to make more overt.)**

We added a paragraph at the end of the results which describes the evolutionary dynamics of the model l396:

“To summarise, social organisation is the equilibrium between two opposing forces, competition within groups where inequality pushes individuals to evolve high influence, and competition between individuals of different groups where efficient group organisation pushes most individuals to evolve low influence. Looking closer, hierarchy provides one direct and one indirect benefit [15] to followers compared to individuals in acephalous groups. First, hierarchy provides a direct benefit to followers because it increases the amount of surplus resources produced and thus, it increases the fitness of followers. Second, hierarchy provides an indirect benefit to followers because it increases the group size and hence the amount of resources produced in the following generation. This increases the fitness of followers’ offspring. The contribution of each benefit is hard to distinguish but their role can be examined by investigating the effect of high migration rate, which suppresses population structure and indirect benefits to offspring on the same patch. Supplementary Figure 5 shows that, considering moderate time constraints, a high migration rate leads to the disappearance of hierarchy at equilibrium. This highlights the importance of the indirect benefit to offspring that remain on the patch in sustaining hierarchy. On the other hand, supplementary Figure 6 shows that hierarchy evolves for any migration rate if the the time constraints are high. In this case, the direct benefit is high enough to overcome the cost of economic inequality. In conclusion, hierarchy can evolve when time constraints are high through the immediate direct benefit of producing extra resources, but the indirect benefit resulting from the feedback loop between hierarchy, group size and scalar stress allows hierarchy to evolve over a much wider range of conditions.”

- **To what extent does the modelled process match what goes on in the real world?**

We have added a justification for the modelled process in the definition of the evolutionary model l194:

“We study the evolution of α using a standard island model with a population of individuals that is subdivided into a finite number of patches, N_p [20]. We consider that group size can vary and thus, groups can compete by differential migration. This process, called demic diffusion, has been observed in humans; for instance in the spread of agriculture. [21]. This assumption is also conservative, as alternatives such as considering that groups can replicate or that group survival depends on group fitness, would increase the competition between groups and favour hierarchy [22].”

- **The only benefit leaders have in this model is their ability to get individuals in a group to coordinate quickly i.e. there is no difference in outcomes or pay-offs for the particular value the group settles upon. What situations in the real-world does this apply to? What differences might we see if there were different outcomes for different values of the trait?**

We added the following sentence to answer this point l230: “ *The consensus time determines the cost of organisation (equation 6), and the outcome of the consensus decision-making determines the share of individuals (equation 7) as explained in the following paragraphs. We do not consider that the decisions taken affect the success of the group except for the distribution of resources. We investigate the emergence of leaders defined as influential individuals, as defined in psychology, e.g. in experiments [23], and as observed in psychological profiles of leaders [12]. In this case where leaders are not better at taking decisions, integrating the effects of the collective decision would only result in more noise and not qualitatively change our results. Exploring the emergence of leaders as more informed individuals can be done in further work but it is not our focus here.*”

- **Related to the last point the applicability and scope of the model would be easier for the reader to understand what situations the model most closely resembles. Rather than being a general model of the evolution or importance of hierarchy it sounds like the authors are modelling something like the emergence of “Big Men” style societies where certain charismatic individuals attempt to influence members of their group. In the literature this often referred to as “achieved” rather than “ascribed” leadership because there is no office of leader that must be filled and the role is not passed on in an hereditary manner.**

In order to clarify the scope of the model, we added the following sentence in the introduction l59:

“Previous work has shown that scalar stress can drive the evolution of institutionalised hierarchy [10], where a leader is appointed by a centralised process. However, rather than being ascribed, hierarchy is likely to initially emerge from the evolution of intrinsic physical and psychological traits of individuals, e.g. height [11], talkativeness and charisma [12]. Emblematic examples of such informal hierarchy are the “big man” societies observed in Polynesia and New Guinea, in which leaders are defined by their persuasion skills rather than by an ascribed position [13].”

- **To fit with the broad scope of PRSB it might be worth saying to what extent this model could apply to non-human animals too and provide examples.**

We thank the reviewer for pointing out that this model might give insights on the evolution of leadership in non-human species. The space available

in this manuscript would not allow to give this discussion the level of detail it requires but we have added the following at the end of conclusion 1505:

“Although the iron law was initially proposed to explain human social organisation in the post industrial revolution era, our model suggests that its explanatory scope might be wider than first believed. For example, there is evidence that even non-human species also use consensus decision-making to coordinate [24]. While it is hard to draw conclusions about other species with the current model, further work could tailor the model to investigate the emergence of leadership in these species.”

- **Lines 13-14 The independent emergence of hierarchy and its pervasive presence suggest an evolutionary origin [4]. Unclear what you mean by evolutionary origin here**

We have modified the sentence to make it more explicit 113:

“The independent transitions to hierarchy and its pervasive presence suggest that the emergence of hierarchy is (at least partly) the result of natural selection [25].”

- **Lines 19-22 “Humans have inherited traits and preferences towards hierarchy from their primate ancestors, who were organised in dominance hierarchies where individuals physically compete for rank, resources and partners [2]” But this runs counter to the egalitarian ethos that hunter-gatherers are well-known for and even western people in lab studies have been shown to have inequity aversion**

The point stressed by the reviewer is explained in the sentences following the citation 120:

“Humans have inherited traits and preferences towards hierarchy from their primate ancestors, who were organised in dominance hierarchies where individuals physically compete for rank, resources and partners [1]. But adaptations such as the capacity to form large coalitions and the development of throwing weapons led early human groups to reverse this hierarchy [26, 27]. In pre-Neolithic tribes, coalitions of followers imposed strong dominance on uprising leaders and successfully avoided coercive leaders for hundreds of thousand of years [1].”

Referee: 3

this is an exceedingly complex model and anything the authors can do to help the reader fit it all together would be appreciated. For

example, it might be helpful to have a single table where all the parameters are defined. The potential parameter space is very large and is only lightly sampled. (I could not see the supplementary materials however as I explained to the editor.)

We have modified the manuscript to address these comments as far as possible. Unfortunately, there is not enough space to add a table with all the parameters but we list the parameters at the beginning of the analysis.

- **I had trouble interpreting several of the key figures. On 1 2, what is nsub1? The population structure has not (yet at least) been explained. One assumes that nsub1 is the size of the first group, but how does that relate to the x axis where group size ranges from 0 to 1000?**

The n_1 that the reviewer discussed is in truth n_l . The confusion is due to the letter “l” being similar than the number “1” in the font used. To avoid this confusion, we spell out the full name of the variables instead of the symbols in the figures.

- **Similar issue with 2; how can groups of 30 have 50 leaders? Also 2, I presume that the linear regression coefficient depicted is the slope (should be explicit).**

We have added the following sentence in the caption of figure 1 (and supplementary figure 1 and 2):

“The group sizes considered for the replicated simulations are from 50 to 1000 with an increment of 50.”

We added the following sentence in the caption of figure 1: *“Scalar stress measured by the linear regression coefficient (slope) of time to reach consensus on group size as a function of number of leaders.”*

- **On 4, I wonder if it might be mislabeled. Is it the mean of alpha or its mean skewness that is indexed by the red/blue scale? Does Mean size by patch really indicate Mean group size by patch? What is meant by “Proportion of individuals in the population”? I think what we’re supposed to see is that through time skewness of alpha goes up, as does mean group size, but I can’t decode it.**

To clarify Figure 4, we changed the title of the y-axis from mean group size by patch to average group size. It is also specified l308:

“The results presented are the average across patches when the result is a function of generations, and the average across patches, generations and simulations when the result is a function of a parameter.”

We also added the following sentence to the caption of Figure 4:

“For a given generation, the colour is composed of sections with each section showing the proportion (size of the section) of individuals with a given

influence (colour of the section).”

- **On Figure 5C, blue, are those resources per capita, or simply in total? It seems to me that it’s the per capita figures that would be critical.**

The blue curve in the Figure 5C represents the total resources. As pointed out by the reviewer, resources by capita would be a useful information but we do not represent the per capita resources here as it differs between individuals within the same group (the resources are not equally distributed between all individuals when $d > 0$).

- **General comment: since leadership has somewhat different effects here (decreasing opportunity costs as group sizes grow, thus allowing group sizes to grow which in turn allows them to enjoy increasing returns to scale) its hard to parse at any point which of these effects dominates. Most likely they are co-constitutive and equally important, but I couldn’t glean that from the ms.**

We added a paragraph at the end of the results which describes the evolutionary dynamics of the model l396:

“To summarise, social organisation is the equilibrium between two opposing forces, competition within groups where inequality pushes individuals to evolve high influence, and competition between individuals of different groups where efficient group organisation pushes most individuals to evolve low influence. Looking closer, hierarchy provides one direct and one indirect benefit [15] to followers compared to individuals in acephalous groups. First, hierarchy provides a direct benefit to followers because it increases the amount of surplus resources produced and thus, it increases the fitness of followers. Second, hierarchy provides an indirect benefit to followers because it increases the group size and hence the amount of resources produced in the following generation. This increases the fitness of followers’ offspring. The contribution of each benefit is hard to distinguish but their role can be examined by investigating the effect of high migration rate, which suppresses population structure and indirect benefits to offspring on the same patch. Supplementary Figure 5 shows that, considering moderate time constraints, a high migration rate leads to the disappearance of hierarchy at equilibrium. This highlights the importance of the indirect benefit to offspring that remain on the patch in sustaining hierarchy. On the other hand, supplementary Figure 6 shows that hierarchy evolves for any migration rate if the the time constraints are high. In this case, the direct benefit is high enough to overcome the cost of economic inequality. In conclusion, hierarchy can evolve when time constraints are high through the immediate direct benefit of producing extra resources, but the indirect benefit resulting from the feedback loop between hierarchy, group

size and scalar stress allows hierarchy to evolve over a much wider range of conditions.”

- **Starting in 287:** **”On the one hand, an efficient hierarchical organisation allows a group to produce a larger amount of resources and hence reach a larger size. On the other hand, hierarchy provides a stronger advantage as group size increases because the cost of organisation increases less in hierarchical groups than in egalitarian groups.”** What (in the model) if anything ever puts the brakes on this process? What (in the reference societies) might do this?

The positive feedback loop between social organisation and production is limited (in the model) by the diminishing returns to scale that happens when the group is above a certain size. This diminishing returns to scale (in real world) is due to a number of factors which limits the productivity of a group, for instance lands available or technological levels. To clarify, we added l390:

“There is a feedback loop between hierarchy leading to larger group size, and larger group size increasing the benefit of hierarchy. Eventually, this feedback loop comes to an end due to diminishing marginal returns i.e. other limiting factors than group size. Yet, this feedback loop amplifies the benefit that hierarchy provides to group members and favours its evolution.”

- **”Economies of scale”** are mentioned a number of times, e.g., **”We make the assumption of economy of scale in which additional participants increase the benefit superlinearly [31].”** I think the more common usage is to speak of increasing returns to scale, and I think that’s what’s actually meant (given the reference to superlinear).

We have replaced the term “economy of scale” by “increasing returns to scale” in the manuscript.

- **There are some missed opportunities to link to other relevant work.** For example, in a very recent issue of *Antiquity*, Amy Bogaard and co-authors show why agriculture where returns are limited by labor is less likely to produce wealth inequality than is agriculture limited by land. (I’m traveling and don’t have access to the exact citation.) This seems (superficially at least) to be at variance with the statement **”The benefit of hierarchy depends on the time constraints C_t , which translates the time to reach consensus into an opportunity cost of organisation. ...**

For a low level of time constraints, the benefit of hierarchy has a negligible effect on organisation and group production and thus, hierarchy does not evolve.” Why this disjunction? Interesting since both of these statements ostensibly target Neolithic societies.

We thank the reviewer for directing us on the interesting work of Amy Bogaard. The reviewer points out a particular publication: [28]. However, we do not think this publication would fit in this manuscript. This is because this publication looks at the evolution from inequality to more inequality rather than from egalitarian to hierarchy. To roughly summarise, the thesis of the authors is that initial inequality can be reinforced by wealthy individuals monopolising expensive technological innovation (in their cases, ox farming) to create surplus resources and to increase their wealth. Then, it allows the wealthy individuals to get more lands. Ultimately, it leads to few wealthy individuals and a majority of subordinate individuals (as in patron-client theories [29]). The model presented in this manuscript differs as it looks at the emergence of hierarchy (and then inequality) in previously egalitarian groups. As discussed in the introduction, these theories struggle to explain the emergence of hierarchy when forms of coercion were absent and inequality was actively avoided by followers. Please note that these explanations are not mutually exclusive but rather, the paper of *Bogaard et al* could explain how the hierarchy and economic inequality present at the end of our simulations becomes exacerbated later on.

Concerning the parallel with their work and our results on time constraints on consensus, we argue that the costs discussed in these two manuscripts are different. The paper of *Bogaard et al* [28] looks at the cost of labour e.g. amount of work required to cultivate a portion of lands, while we look at the cost of group organisation e.g. cost of agreeing on common rules to share water between farmers.

- **Are there any empirical data available that could inform us as to whether the levels of alpha used in the model for informal hierarchy are realistic? If so, please cite.**

It is not clear for us if the reviewer means (i) the range of value on which α is defined or (ii) the values of α observed at equilibrium in the evolutionary model. In the first case, as far as our knowledge goes, there is not data which could directly inform us on the range of possible value of influence. We agree that these data could be useful, but in the absence of these data, our model explores a wide range. In the second case, the level of abstraction required for this kind of mechanistic model does not aim to make precise quantitative predictions, but rather predicts qualitative

trends.

- **Also, quite a few uncited works come to generically similar conclusions about group size and leadership. Even though these do not (so far as I can recall) contain opinion formation models, it may be that some of these suggest that such a model is not strictly necessary? For example: Crabtree, Stefani A., R. Kyle Bocinsky, Paul L. Hooper, Susan C. Ryan, and Timothy A. Kohler 2017 How to Make a Polity (in the central Mesa Verde region). *American Antiquity* 82(1):71-95. (See also references therein.)**

We thank the reviewer for pointing out these publications. These are important work in the field but we would argue that there are differences between these prior models and the one presented in the manuscript. First, Crabtree et al. look at the evolution of nested hierarchy, in the sense of polity inside polity. We look here at the evolution of distribution of influence.

Second, the leaders in Crabtree et al. play the role of punishing the defectors. Although we acknowledge the importance of this explanation (as done in the discussion), the paper here presents another role of leaders, which is to facilitate coordination.

Third, Crabtree et al. are looking at a particular system (Pueblos populations) and so their model is more detailed than our abstract model. As a result, they integrate a large number of mechanisms including mechanisms that could lead to the emergence of hierarchy, such as demographic increase, conflict between groups with group dominating other, and possible revolution of subordinates against dominant. While this provides a more realistic simulation to understand the evolution of polities in a particular set of societies, our abstract and mechanistic model highlights a more general mechanism. For instance, we provide a detailed account of the forces acting on the evolution of hierarchy and demonstrate clearly the role of scalar stress.

- **Also, you cite a working paper that has now been published, as follows: Hooper, Paul, Eric Smith, Henry Wright, Timothy Kohler, and Hillard Kaplan 2018 Ecological and Social Dynamics of Territoriality and Hierarchy Formation. In *The Emergence of Pre-Modern States: New Perspectives on the Development of Complex Societies*, edited by Jeremy Sabloff and Paula L. W. Sabloff, pp. 105-132. Santa Fe Institute Press, Santa Fe.**

We thank the reviewer for pointing out this published version. However, we replaced this citation by the citation of an earlier paper from these authors, paper which demonstrates the evolution of hierarchy with leaders

as monitors (Hooper et al, 2010): [18].

- **Trivia: typos of some sort in lines 107, 444; Ref 12 incomplete.**

We fixed the typos and we removed the reference 12.

References

- [1] Boehm C. Hierarchy in the forest: The evolution of egalitarian behavior. Cambridge: Harvard University Press; 2001.
- [2] Glowacki L, Von Rueden C. Leadership solves collective action problems in small-scale societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2015;370(1683):20150010.
- [3] Bass BM, Riggio RE. Transformational leadership. Psychology Press; 2005.
- [4] Michels R. Political parties: A sociological study of the oligarchical tendencies of modern democracy. New York: The Free Press; 1911.
- [5] Kearns M, Judd S, Tan J, Wortman J. Behavioral experiments on biased voting in networks. *Proceedings of the National Academy of Sciences*. 2009 2;106(5):1347–52.
- [6] Hare AP. A Study of interaction and consensus in different sized groups. *American Sociological Review*. 1952;17(3):261.
- [7] Manners GE. Another look at group size, group problem solving, and member consensus. *Academy of Management Journal*. 1975;18(4):715–724.
- [8] Alberti G. Modeling group size and scalar stress by logistic regression from an archaeological perspective. *PLOS One*. 2014 3;9(3):e91510.
- [9] Johnson GA. Organizational structure and scalar stress. In: *Theory and Explanation in Archaeology*. New York: Academic Press; 1982. p. 389–421.
- [10] Powers ST, Lehmann L. An evolutionary model explaining the Neolithic transition from egalitarianism to leadership and despotism. *Proceedings of the Royal Society B: Biological Sciences*. 2014;281(1791).
- [11] Berggren N, Jordahl H, Poutvaara P. The looks of a winner: Beauty, gender and electoral success. *Journal of public economics*. 2010;94(1-2).
- [12] Judge TA, Bono JE, Ilies R, Gerhardt MW. Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*. 2002;87(4):765–780.

- [13] Sahlins M. Poor man, rich man, big-man, chief: Political types in Melanesia and Polynesia. *Comparative Studies in Society and History*. 1963;5(3):285–303.
- [14] Von Rueden C, Van Vugt M. Leadership in small-scale societies: Some implications for theory, research, and practice. 2015;.
- [15] Frank SA. *Foundations of social evolution*. Princeton University Press; 1998.
- [16] Makowsky MD, Smaldino PE. The evolution of power and the divergence of cooperative norms. *Journal of Economic Behavior & Organization*. 2016 6;126:75–88. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S016726811500236X>.
- [17] Cavalli-Sforza LL, Feldman MW. *Cultural transmission and evolution: A quantitative approach*. vol. 16. Princeton: Princeton University Press; 1981.
- [18] Hooper PL, Kaplan HS, Boone JL. A theory of leadership in human cooperative groups. *Journal of Theoretical Biology*. 2010;265(4):633–646.
- [19] Ostrom E. *Governing the commons*. Cambridge University Press; 1990.
- [20] Wright S. Evolution in mendelian populations. *Genetics*. 1931 3;16(2):97–159.
- [21] Pinhasi R, Fort J, Ammerman AJ. Tracing the origin and spread of agriculture in Europe. *PLOS Biology*. 2005;3(12):1–9.
- [22] Lehmann L. The demographic benefits of belligerence and bravery: Defeated group repopulation or victorious group size expansion? *PLoS ONE*. 2011;6(7).
- [23] Sorrentino RM, Boutillier RG. The effect of quantity and quality of verbal interaction on ratings of leadership ability. *Journal of Experimental Social Psychology*. 1975 9;11(5):403–411.
- [24] Conradt L, Roper TJ. Consensus decision making in animals. *Trends in Ecology & Evolution*. 2005 8;20(8):449–456.
- [25] Van Vugt M, Hogan R, Kaiser RB. Leadership, followership, and evolution: Some lessons from the past. *American Psychologist*. 2008;63(3):182–196.
- [26] Bingham PM. *Human uniqueness: A general theory*. University of Chicago Press; 1999.
- [27] Schaik CPV. *The primate origins of human nature*. Wiley-Blackwell; 2015.
- [28] Bogaard A, Fochesato M, Bowles S. The farming-inequality nexus: new insights from ancient Western Eurasia. *Antiquity*. 2019;93(371):1129–1143.

- [29] Eisenstadt SN, Roniger L. Patron-client relations as a model of structuring social exchange. *Comparative Studies in Society and History*. 1980 1;22(1):42-77. Available from: https://www.cambridge.org/core/product/identifier/S0010417500009154/type/journal_article

Appendix B

Response to editor and reviewers - Minor revision From disorganised equality to efficient hierarchy: How group size drives the evolution of hierarchy in human societies

We thank very much the editor and the reviewers for their positive response and their additional comments. Please find below our response to the editor and reviewers' comments, which includes the list of modifications done to the manuscript.

The reviewers' comments are shown in bold, citations of the manuscript in italic and changes to the manuscript are shown in italic and bold. We show only the comments which required a response.

1 Editor

- **Reviewer 2 also suggests making the code available, which is something I would strongly encourage the authors to consider. Although in theory all simulations are reproducible thanks to clear and detailed descriptions, access to the code would allow the interested reader to learn more about the practical implementation.**

The code has been made available online and a data availability statement has been added to the manuscript.

- **The papers mentions the "Big Man" style of leadership in "New Guinea and Polynesia" - New Guinea is OK, but Polynesia is more often characterized by formal, institutional forms of "chiefly" leadership. I suggest using the term "Melanesia" to refer to New Guinea and the surrounding area, as this is where "Big Man" style leadership is most closely associated**

We thank the reviewer for the clarification. We replaced the terms "New Guinea and Polynesia" by "Melanesia".

- On line 181 suggest rather than "economic inequality" use the term "inequality in resources" (or similar) to make things more general as readers may think of "economic" in narrow terms

We replaced the term "economic inequality" by "inequality in resources" in the manuscript.

- lines 196-198 I appreciate that the authors have tried to address my comments about the relevance of the model to real-world processes, but I found the references to "demic diffusion" and competing by differential migration confusing in the context of this model. My comments about applicability of the model relate mainly to thinking about examples from the real-world where decision-making processes and contexts in which reaching consensus are a) important, and b) well described by the mechanism explored in this model. "Demic diffusion" is usually discussed in situations of population *expansions* where by a trait increases population growth and populations begin to spread from the initial source of innovation. I would normally associate it with a large-scale spatial trend in the spread of a trait - which does not seem to be the case being modelled here. The authors rightly note that demic diffusion has been linked to the spread of agriculture but the model is not about the spread of agriculture. So even if in some sense the model could be described as representing demic diffusion, I think this may confuse readers.

Thank you for clearing up this point. We misunderstood previous comment. We removed the following paragraph:

"This process, called demic diffusion, has been observed in humans; for instance in the spread of agriculture. [1]. This assumption is also conservative, as alternatives such as considering that groups can replicate or that group survival depends on group fitness, would increase the competition between groups and favour hierarchy [2]."

To answer the point a), we added the following sentences:

*"The iron law of oligarchy [3] and behavioural experiments [4] suggest that the benefit of hierarchy on group coordination lies in its effect on the time a group spends to reach consensus and take a collective decision. **Consensus decision-making is an efficient method for a group to coordinate, in particular to tackle tasks where learning the optimal strategy by trial and error is too costly. Examples of consensus decision-making include tribe gatherings to discuss the next camp location, councils of war to decide upcoming battle strate-***

gies, or parliamentary debates on new laws. Yet, the time spent to reach consensus (consensus time in short) is costly because individuals dedicate time to organisation instead of carrying out the actual task, and because time itself can carry a cost, e.g. resources get depleted. Thus, we explicitly model the consensus decision-making and the effect of hierarchy on this process. ”

To answer the point b), we added the following sentence with references in support of decision-making as described in our model:

Using this model, we aim to answer the following question: “does hierarchy limit the effect of scalar stress, and if yes, could it drive the evolution of leader and follower behaviours even if it creates inequality in resources? To do so, we consider that in absence of advanced institutions such as voting systems, collective decision-making is a sequence of communications, as observed in human groups faced to coordination problems in laboratory experiments [5] or in real world [6, 7]. Thus, we mathematically describe collective decision making by an opinion formation model, which consists of a sequence of discussions between individuals until a global consensus is reached [8]. ”

- **I understand that constraints of space make it difficult to include the schematic diagram in the main paper but I think this would still be useful in the supplementary materials.**

A diagram for each model has been added to the supplementary material (Figure 1). We added the following sentences in the main manuscript at the beginning of the definition of each model:

“The outline of the model is represented in the top part of supplementary Figure 1.”

“The life cycle of the evolutionary model is represented in the top part of supplementary Figure 1.”

In case, the reviewer can not access the supplementary materials, we added a copy of this figure at the end of this document.

References

- [1] Pinhasi R, Fort J, Ammerman AJ. Tracing the origin and spread of agriculture in Europe. PLOS Biology. 2005;3(12):1–9.
- [2] Lehmann L. The demographic benefits of belligerence and bravery: De-feated group repopulation or victorious group size expansion? PLOS ONE. 2011;6(7).

- [3] Michels R. Political parties: A sociological study of the oligarchical tendencies of modern democracy. New York: The Free Press; 1911.
- [4] Kearns M, Judd S, Tan J, Wortman J. Behavioral experiments on biased voting in networks. *Proceedings of the National Academy of Sciences*. 2009 2;106(5):1347–52.
- [5] Cooper R, Dejong DV, Forsythe R, Ross TW. Communication in the battle of the sexes game : Some experimental results. *The RAND Journal of Economics*. 1989;20(4):568–587.
- [6] Von Rueden C, Gurven M, Kaplan H, Stieglitz J. Leadership in an egalitarian society. *Human Nature*. 2014;25(4):538–566.
- [7] Boehm C. *Hierarchy in the forest: The evolution of egalitarian behavior*. Cambridge: Harvard University Press; 2001.
- [8] Castellano C, Fortunato S, Loreto V. Statistical physics of social dynamics. *Reviews of Modern Physics*. 2009 10;81(2):591–646.

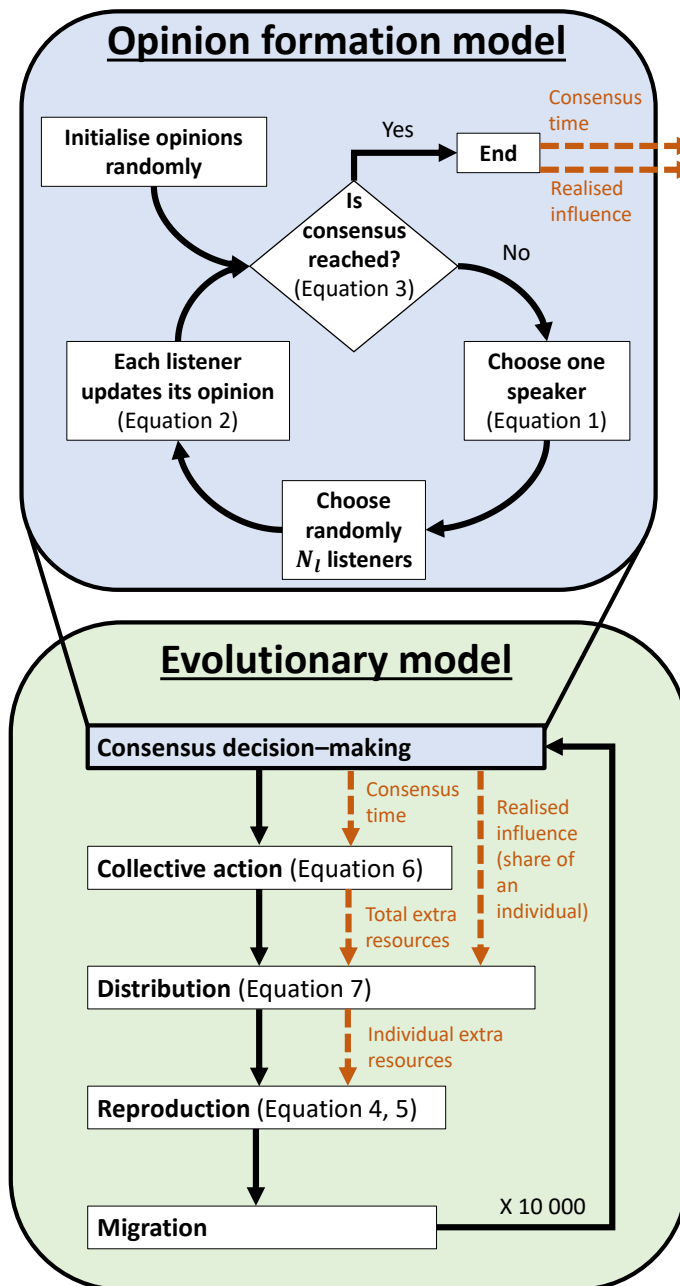


Figure 1: Diagrams representing the opinion formation model (top) and the life cycle of the evolutionary model (bottom). A box represents a process. A diamond represents a condition. The black arrows show the next step. The dashed orange arrows represent the important variables that feed in a process or result from a process. For instance, the distribution of resources depends (i) of the realised influence of an individual and the total extra resources, and (ii) provides the amount of individual extra resources. The consensus decision-making in the life cycle of the evolutionary model is described by the opinion formation model represented on the top. The fitness is calculated in the reproduction step and strongly depends of the extra resources received by an individual.