Editor  
Comments to Author:  
The board member notes that the terse style of the presentation is a bit of a departure from the more common style in the journal. While I am comfortable with the way it is now, the authors may wish to alter the presentation for greater impact on the typical audience of this journal.

TODO: YR  
  
Associate Editor  
Board Member: 1  
Comments to Author:  
This manuscript presents a model for the evolution of cooperation when there is both vertical and nonvertical cultural transmission of the trait. The model also allows for different degrees of horizontal transmission between individuals who are undergoing fitness determining social interaction and, interestingly, it also allows this “interaction-transmission” association probability to evolve. I think the manuscript does an excellent job of providing some synthetic results for the general issue of the evolution of cooperation and the findings will likely be of interest to many of the readers of Proc B. Both referees were also strongly in favour of publication, although only one of these provided substantive comments. I agree with all of their suggestions and have two further comments of my own.

- We thank the editor…. TODO  
  
First, I think it is important to have a very clear description of the structure of the model and Figure 1 partly does this job. I wonder, though, if it can somehow incorporate the ordering of the different events in the life cycle more clearly as well.

TODO: DC, change order panels d->c->b and then a. fix pixelized people in c.

Related to this, L87 refers to Fig 1a when it should be Fig 1d.

- Fixed  
  
Second, although I appreciate the rather terse, point-form sort of presentation in the manuscript, it does seem like a bit of a departure from the usual style of Proc B papers (which tend to have a more narrative structure). Indeed, it reads like a paper one might find in a more technical journal. This is an editorial issue though so I will leave it to the editor to decide if there is anything that should be done about this.  
TODO: YR  
  
  
Reviewer(s)' Comments to Author:  
  
Referee: 1  
  
Comments to the Author(s)  
In “Non-Vertical Cultural Transmission, Assortment, and the Evolution of Cooperation,” the authors present a discrete-time dynamical system for the frequency of a cooperative trait in the presence of both social and genetic transmission. After defining the model and its parameters, the authors summarize its equilibria and their stability, presenting several conditions, in the form of inequalities, characterizing the various outcomes (fixation of one type or polymorphism). The social transmission parameter is then allowed to evolve via a modifier locus, and it is shown that the resulting dynamics favor lower population fitness.  
  
Overall, I enjoyed reading the paper and I’m supportive of it being published in Proceedings B. One disadvantage of the model is that it has many parameters, but at the same time it seems to be about as simple as possible while capturing all of the qualitative features of interest. I also liked that the analysis done in the appendix is somewhat clean, and a relatively complete description of the evolutionary dynamics can be described in the main text.

- We thank the reviewer… TODO

I do, however, have a number of specific comments, which can be found below.  
  
Comments:  
  
-References should be ordered by appearance if using bracketed numbers.

- Fixed  
  
-The literature surrounding equations 1-4 is just a small portion of work related to conditions of this form. One obviously cannot cover everything, but studies of the effect of space on altruism often result in conditions of the form b > {something} \* c. These include approaches using both inclusive fitness (e.g. “Evolution of cooperation in a finite homogeneous graph” by Taylor, Day, and Wild) as well as direct analyses of the spatial structure (e.g. “Evolutionary dynamics on any population structure” by Allen et al.). The text would benefit from some further discussion of these kinds of results (of the authors choosing—the above are just some examples).

TODO: YR  
  
-Around line 64, I find it odd to say that $\varphi$ takes the role of relatedness. I understand that this is probably meant to compare the structural form of this inequality to Hamilton’s rule, but too often it seems that people try to suggest that something should be interpreted in terms of relatedness when it is not appropriate to do so. I would remove this sentence to prevent misunderstanding.

~~TODO: DC~~

- We changed line 61 from “phi takes the role of relatedness (…) or assortment (…)” to “phi replaces relatedness (…) or assortment (…)” to avoid incorrect interpretation.  
  
-At the end of line 82, the authors say that evolution of cooperation can be enhanced “partly because [cultural transmission] can diminish the effect of natural selection.” Later on in the paper, the meaning of this claim becomes clearer, but it really left me scratching my head as to what it could mean when stated so briefly at the end of the introduction. It would be helpful to have a slightly more detailed explanation of what to anticipate from this statement.

TODO: YR  
  
-Around line 92, $\tilde{p}$ looks nearly identical to $\hat{p}$ unless one looks closely. Can the authors change these symbols so that they appear more distinct in the subsequent equations?

- To avoid confusion, we changed $\tilde{p}$ to $\dot{p}$.  
  
-Equation 8 needs to be described more clearly, especially if the paper is to be published in Proceedings B. A careful reading suggests that individuals interact in pairs and obtain payoffs (fitness) from these interactions. These payoffs are to be interpreted in a relative sense, with competition happening proportional to payoff (globally). There is then reproduction based on this competition, and the offspring inherits a trait based on horizontal transmission. Vertical transmission, which is described on the previous page, is then taken into account, which together with equation 8 gives a complete description of the dynamical system, equation 9. The latter (about vertical transmission) is briefly mentioned following equation 8, but I really think that the paper would benefit from several sentences—or even a paragraph—completely explaining all of the terms that appear in equation 8 and the underlying assumptions (e.g. infinite panmictic population versus sampling with stochasticity, etc.).

TODO: YR

Line 84 – “very large well-mixed population”

Line 110, remove parentheses around “i.e. after selection”

Explain eq 8  
  
-In “Result 1” (line 130), “a cooperation” should be “cooperation.”

- Fixed  
  
-Please discuss conditions 1-4 (lines 132-136) in more detail, as it relates to the mathematical properties of equilibria.

TODO: YR

(Might also be worthwhile to mention that “stable polymorphism” is sometimes called “coexistence” and “unstable polymorphism” is “bistable competition”).  
~~TODO: DC, add to line 137: “that “stable polymorphism” is sometimes called “coexistence” and “unstable polymorphism” is “bistable competition.”~~

- We mentioned this around line 139.

-On line 144, by “unites” you just mean that the two conditions predict the same selection pressures locally, right?

~~TODO: DC, change “unite” to “merge”. Add remark here that we mean that if we merge these conditions, we get eq 12.~~

- No, by unities we meant that if we merge the conditions for fixation of cooperation and coexistence (stable polymorphism) we get Eq12.

-Line 160: “…this condition cannot be formulated in the form of Hamilton’s rule due to the bias in horizontal transmission…” This is an important point, as Hamilton’s rule applies to a rather special class of models. I wonder, though, what is the purpose of this statement here? I am not suggesting that the authors remove it, but rather that they explain why one might expect a version of Hamilton’s rule in this case.

~~TODO :DC~~

- To address the reviewer’s comment, we replaced in line 160 “In general” with “Importantly” and we changed “in the commonly used form of Hamilton’s rule” to reflect that Hamilton’s rule is commonly used in the literature and textbooks. This also corresponds to the manuscript’s introduction.  
  
-Later on, around line 166, the authors state that the quantity $\alpha T/(1-T)$ “can be regarded as the effective relatedness” in Hamilton’s rule. I don’t like this interpretation, as it seems to try to shoehorn something cultural into a biological statement. Sure, there are elements of both in the model, and this particular case pertains to the case $T\_A = T\_B$, but I still see very little value in describing this as an “effective relatedness,” and I think it could create confusion where there ought to be none.

~~TODO: DC change “effective” to “cultural”~~

- To avoid confusion, per the reviewer’s comment, we replaced “effective” with “cultural”: “PASTE NEW SENTENCE HERE”. We consider “cultural relatedness” to reflect the probability that two individuals have the same cultural trait, much like genetic relatedness is the probability that two individuals have the same genetic trait.

-Line 192, please write “sign()” in words in the main text, e.g. “the sign of […] coincides with the sign of […].” It is fine to keep this notation in the appendix (as it is), though.

~~TODO: DC – write reply here.~~- Fixed  
-Line 228: $\alpha$ must have been written as $alpha$ since the symbol is not rendered.

~~TODO DC~~

- Fixed

-Line 242: The paragraph ending here is quite interesting, and I wonder whether a bit more intuition for the detrimental effects of the evolution of \alpha can be provided here?

TODO YR  
  
-In the section on population structure, the model setup had me a bit confused. Half of the population is chosen to initiate interactions, but are these really interactions or are they just actions? The game being considered is completely additive, which means a two-player interaction can be decomposed into two actions, one from each player. This confusion is compounded by the later expression for fitness as $w=1+b\*n\_b-c\*n\_c$, where both $n\_b$ and $n\_c$ are interaction counts. So if X and Y are neighbors and are both chosen to initiate an interaction, what happens if X chooses Y but Y chooses someone else? Does X still get a benefit from Y, or are these expressions directional (as they seem like they should be)?

TODO DC  
- Indeed half of the population is chosen to initiate interactions, each with a randomly chosen neighbor.

Thus the expected number of interactions each host participates in is 1.

But the number of interactions in which a host can participate may be 0 (if the host was not chosen to initiate interaction and it was not picked by any of his neighbors for interaction)

or 1 or greater than 1 (if the host was chosen to initiate an interaction, and it was also picked by a neighbor for an interaction).

A note: We model this procedure in this way (of half of the population initiating interactions with theri neighbors) in order to overcome the problem of how to partition a population on a lattice into couples.

Moreover, a host that participates in more than one interaction can change its behavior from one interaction to the other, due to horizontal transmission.

This is why we calculate for each host in how many interactions it cooperated (not necessarily 0 or the number of interactions) and in how many interactions it received help.

-Line 254: a “grid” seems like a lattice with a von Neumann neighborhood, but here the authors refer to a Moore neighborhood ($M=8$); is this correct? Maybe mention this up front as well.

~~TODO DC~~ Still need to find REF for Moore neighborhood.

- We indeed refer to a Moore neighborhood. We now mention this in line XXX: “On an infinite grid, M=8 (i.e., Moore neighborhood REF),”  
  
-Line 274: “…under complex scenarios.” While it is true that the population is structured, in some ways it is nearly unstructured since it is so homogeneous. “Complex scenarios” to me would indicate something more topologically heterogeneous. I don’t think that the authors need to consider anything more complex here, but I would at least comment on the homogeneity of the primary structured population in the paper. A similar comment applies to lines 300-302 in the discussion section.  
TODO YR

-In line 324, the authors claim that this mechanism does not require population structure. However, it feels like there is a version of population structure baked into the model via its parameters. I think I understand what is meant here, but some minor clarification would be nice.

~~TODO DC~~

- To clarify our discussion, we changed “population structure” to “spatial structure”.

-In the conclusion around line 336, what is being referred to is that the model effectively generates identity by state that does not come from identity by descent; is this correct?

- This is correct, but because transmission in our model is cultural, we do not use the term “by descent”.  
  
-Figure 1: “showing The fitness” should be “showing the fitness”. Also, I believe that “prisoners’ dilemma” should be “prisoner’s dilemma” because the dilemma applies to the individual and not the group (though there is a conflict of interest between the two).

~~TODO: DC~~

- Fixed in Fig 1 and in line 253.  
  
Referee: 2  
  
Comments to the Author(s)  
The paper studies the evolution of cooperation and interaction-transmission association using simulations and analytical methods. The study is carefully carried out and the manuscript is well-written.

- We thank the reviewer for his considerate evaluation.  
  
I have one question about their choice of the payoff matrix. In fact, they use a simplified Prisoner's Dilemma -- so-called donation games, to model social interactions. How about using a general payoff matrix with R, S, T, P?

TODO: DC verify that this is correct.

we focus on c<b (donation game – HILBE et al. Evolution of extortion in iterated prisoner’s dilemma games. Proceedings of the National Academy of Sciences, 2013, 110.17: 6913-6918.‏), as we are interested in the case where cooperation is advantageous to the population but disadvantageous for the individual.  
  
Also vertical transmission is not explicitly modeled using diploid sexual reproduction with recombination/mutation. How their results under simple assumption would change under more realistic vertical transmission?

~~TODO: DC~~

In our paper we consider cultural transmission rather than genetic transmission, and, therefore, there is no concept of diploids or sexual reproduction. To emphasize this, we changed in line 85 “an offspring inherits its phenotype from its parent via vertical transmission” to “… via cultural vertical transmission.” We assume vertical cultural transmission is uni-parental (stated in line 88). One could assume bi-parental transmission in later work, as in ﻿Cavalli-Sforza & Feldman, 1981.