

Date: August 2, 2022

Thesis report: The influence of social processes and structures on cultural evolution

Candidate: Michael Chimento

Reader: Robert Boyd

Grade: 0

Overall: This is great dissertation. It combines interesting and novel modeling work with a number of really beautiful experiments. All published, I believe, in very high-quality journals. Chimento's command of the literature is extremely impressive. It is one of the best two or three best dissertations that I have read. I grade it excellent = 0.

In what follows I offer comments that Chimento might find useful.

Ch 1

Such a lovely experiment---clearly laid out with very nice results.

Ch 2

The chapter is not very clear about the goals of the exercise, and especially unclear about the relationship between the structure of the model and the underlying evolutionary ("diffusion") processes. It may be that my confusion is due to not keeping up with previous work, but if it were mine I would I think it needs a serious rewrite.

More detailed comments about the model and my confusions about it keyed to page/paragraph or eqn number.

41/3-42/4 $N = 24$ is a very, very small population. In standard population genetics models drift would determine everything unless selection is very strong. You say top of p 42 that there are no drift-like effects. I'm not sure I understand why, but my guess is that it results from a crucial, unjustified, and to me quite odd, assumption which is expressed in the form of Eq 21. Here the probability of acquiring knowledge of the new behavior depends on T_i which in turn depends on the sum of the fraction of times each individual who is connected produced the behavior in the last m time periods. But later you say that the network is fully connected which I think means all of the $a_{ij} = 1$, but in any case they are fixed for the entire time period. If so, all individuals experience the same value of T_i (because from eqn 2.2 when $a_{ij} = 1$, they include their own behavior in computing T_i). There is no sampling, and no drift. But

this is really weird, no? There is no “transmission”, just a bunch of individuals independently learning about the new behavior with some probability. I guess this might be true, but it’s not obvious and really different from lots of CT models. You need to explain why this model differs so strongly from the standard Fisher-Wright model, and why you think it is a better description of the phenomena of interest.

With a fully connected network this model is similar in some ways to the core models in Lumsden and Wilson (1981). They separate transmission and production initially, and then assume that everyone is connected in the population and the key is to understand production and they include frequency and content dependent production rules. The book foundered for other reasons, but and I always thought that separation of knowing and producing was interesting. Pete and I discussed the separation of transmission and production at several points in our 1985 book, especially in the section on how cultural variants compete.

44/3 (eq 2.4) If individual i produced behavior k in period $t-1$, $\pi_k = 0$, otherwise $\pi_k = 1$. This means that if an individual does not produce behavior k in period $t-1$, the expected payoff in period t decreases, and if $p \approx 1$ as it is in some of the simulations, it drops to zero.

I don’t understand this assumption. Agents remember the behavior of N other individuals for m time periods, but can’t remember the experienced payoff from two periods ago? Or am I missing something? You need to explain. This assumption also means that the expected payoff of the two behaviors will be the same at equilibrium, no? Most readers will come to the paper with the assumption that the innovation, b , will have higher payoff, which could have a big effect on your results, no? Again, you need to be clear.

45/1 No sampling in production either (eqn 2.6)

46/- If there were fitness differences between alternative behaviors it would be interesting to compute selection gradients on parameter values.

51/- So the vertical axis is knowledge. Am I right that if you plotted behavior it would asymptote to 0.5? So, what is going on is that individuals are finding out about a new behavior that is no better than the alternative, but because they choose randomly the new behavior spreads? I can see that this might be a useful null model for interpreting experimental results, but it’s not very interesting as a model of cultural evolution more generally and I don’t find it very plausible that these results are insensitive to the neutrality assumption (assuming I got that right).

More generally, it is important to explain the qualitative forces in your model in words. You do a lot of simulations (a good thing) but this paper (and subsequent chapters except 7) are weak on conceptual explanation. In my opinion, doing a good job on this is the key to getting your work known outside of a small circle of cognoscenti.

Ch 3.

This one is more interesting that the previous---the spread of a better/more attractive behavior is the key to understanding cultural evolution as an adaptive phenomenon.

75/2 I think your usage of “cultural selection” is wrong headed. The selective retention of higher payoff behavior is not selection. Selection is a culling process and so depends on the variance in the population. If everybody is the same, there can be no selection. The selective retention of higher payoff behaviors (or the selective production) is like biased mutation. It acts most strongly when everybody

has the lower payoff behavior. Pete and I called it guided variation. I'm not wedded to this terminology, but it is really important to distinguish the very different population dynamics that result. You cite Henrich's diffusion of innovations paper in support of this definition, but the distinction is absolutely crucial there. Later in the paper you compound the confusion by referencing turn-over. ALL intergenerational models of CE have to introduce naïve individuals.

75/3 What you say here is correct, I think, but there is a good reason for it, at least in the case of humans. The big problem for models of human CT (and other longer-lived creatures) is that age structure is crucial, but very hard to model analytically. In the village, kids are born and spend a lot of time around their parents. As they get older they are out running around the village learning from a wider range of people including older kids. Then in their teens they may apprentice themselves to a knowledgeable adult and they have lots of interactions with people from other villages. Michelle Kline, Cristina Moya, Joe Henrich and I have done empirical work on this, as have the BaYaka people like Deniz Salali. I spent a long summer (in 1982) trying to adapt models of genetic evolution in age structured populations, but never got very far. Those models depend on genetic evolution being slow so that populations are in stable age distribution, and that all of the genes get transmitted into a single age class. In human cultural evolution people learn in and from all age classes. This is already a lot harder, but I was able to analyze models in which this occurred, but only if cultural parents were sampled randomly from age classes with no correlation between time periods. But this is obviously stupid. The identity of models in the 0-5 age class will be highly correlated with those in the 5-10 age class. Maybe it's time to use the ABM approach like you do. But for people, the structure of the social networks must change with age.

77/1 Age?

80/2 What is the assumed network structure? Again the populations would be way too small for many plausible models of CT.

86/1 This is a very great tit-centric picture. You seem to be making a general claim, but what is "turnover" in human (or baboon/capuchin/vervet) populations. Is it between group migration?

The second part of this paragraph gives a very clear qualitative description of what's going on in the model. You need to do more of this elsewhere.

86/2 More great tit-centrism. Naïve individuals come into human populations as 3-5 year olds. They aren't doing much sampling for many important traits. By the time they start sampling, they aren't naïve.

Ch 4

Another great experiment. Very, very cool result.

96/2 You've don't have the solutions to Roger's Paradox exactly right. The key to "solving" the paradox is not that animals don't always socially learn. Rather there are three (at least) solutions, (1) individuals socially learn *contingent* on the cost/accuracy of individual cues, (2) small improvements are less costly than large ones and they avoid trying to make large improvements, and (3) most important, I guess, they socially learn from higher payoff individuals.

102/3 Won't dropping the non-spread group affect estimates? Can you calculate the probability of getting lack of spread given the parameter estimates from the other 17? If this is really low, it seems to me you have an estimation problem.

106/2 This is a great result, but it's not clear what exactly to make of it. Selection should favor more social learning when the right cues indicate that the new environment is different from the previous one. There are lots and lots of possible cues in the real world. Humans (and other primates) move to a new group at sexual maturity. Surely they know that they are in a new group. Immigrant vervets to the Cheney/Seyfarth study pops very rapidly learned not to freak out when C&S showed up. Which cue should be important?

Ch 5

This chapter is a useful antidote to strong claims by Migliano and colleagues about network structure, but I think you make way too much out of the results.

123/- Weird to plot \log_{10} time on the x-axis. Again 64 is way too small for any human population where, at a minimum it should be 1000 or so, and that neglects contact between different ELU's.

More seriously, I would interpret these graphs as saying network structure is not important. Your model is built on many artificial assumption. This is fine—it's true of all models---but there are many processes not included that could have much bigger effect than network structure.

124/2 The description of the model is way too terse.

127/- Again, all the same.

Ch 6

Another great experiment. The birds have a little CCE and that's very interesting. But they are really bad at it. Simple problems, everything set up for them, and still most of them don't get it. It would be really interesting to know, mechanistically, why.

Ch 7

An interesting discussion with some problems.

176/1 Whether other animals have CCE is, in my view, not an interesting question. It's about words and definitions, not about the world. It's a follow up to an equally uninteresting question, do animals have culture? Humans are outliers in a bunch of dimensions, but so are birds. The right approach is to try to generate causal explanations for the variation among and within species and then do empirical tests. Don't get caught up in definitional disputes.

177/1-2 Ratchet is not a good metaphor because they only go in one direction. CCE can go in lots of directions---increasing fitness or payoff or the reverse. Lots of CCE is driven by cognitive and affective attractors, and much is related to conventions and other traits with many, many equilibria.

179/2 But the percentage increases are small, no? Magnitudes matter. Human's are utterly dependent on the products of CCE, other animals not so much.

180/2 Cultural fitness is not a useful concept because there are many different processes shaping CCE and some are not selective. We need to work out how it works and then build a terminology that comes from causal models. Your own work distinguishing transmission and provides an excellent counter example---CE is different from genetic/Darwinian evolution and trying to shoe horn it into a Darwinian box just leads to confusion.

182/2 When you say “from the perspective of behavior” it’s like when people talk in a Dawkinsesque way about from the genes perspective. This is fine, but you could be clearer. I think your approach here is very similar to the meme’s eye view of cultural evolution which has been endlessly argued about. You are properly catholic, but you need to relate these ideas to the tons of previous work.

184/2 Reducing the probability of extinction is a group-level benefit. I’m happy with selection acting on cultural groups/norms but many are not.

189/1 Stone tools are not a very good example---there has been a lot of work on production. Knapping is hard to learn and it takes a lot practice to get good at it. There has been much less work on the CT and CE of *design*. So, should points be tanged? Can I improve the knapping quality of stone by heating? Should I add ochre to gum adhesive used to haft points? The answers to these questions can be very hard for individuals to learn on their own, but very easy to learn socially. Everybody buries the silcrete under the fire, makes tools w/o tangs, and adds ochre to the adhesive. You do the same. Piece of cake. No long practice/learning period.

Now, skill learning is important for lots of human traits. You learn that using stone tools is very important and the are made using, let’s say, soft percussion which involves a bunch of easy to learn design items, and a bunch of really, really hard to learn skills. But without the easy to learn stuff you’d never get there. And there are many, many design questions that are essential to surviving in every human environment.

191/2 I’m a big Simon Kirby fan but you misrepresent the state of play in language evolution. People in that field don’t agree about anything.

Sincerely yours

Robert Boyd

A handwritten signature in black ink, appearing to read 'Robt Boyd', with a stylized, cursive script.

Professor