Weak content preferences stabilize culture

Alberto Acerbi

Benoît de Courson

Abstract

Outcomes in the cultural arena are due to many factors, but are there general rules that can suggest what makes some cultural traits successful and others not? Research in cultural evolution theory distinguishes factors related to social influence (such as copying from the majority, or from certain individuals) from factors related to intrinsic features of cultural traits (such as being more effective, easy to transmit, or memorable). Here we show, using analytical and individual-based models, that preferences for content, even when weak, being stable and directional, determine the equilibrium point of cultural dynamics when acting together with non-directional social influence. The results have implications regarding the importance of keeping into account individual-level, non-social, factors, when studying cultural evolution, as well as regarding the interpretation of cross-cultural regularities, that have to be expected, but can be product of weak directional forces, intensified by social influence. In addition, they suggest that when planning policies for behavioral changes, it is fundamental to consider widespread individual preferences, even subtle ones, and design interventions that reinforce them.

Introduction

On the first Thursday of March, the UK and Ireland celebrate World Book Day, and in many primary schools children dress up as their favorite book characters. When children turn seven or eight, parents start to recognise in the courtyard the familiar figures of the Hogwarts School of Witchcraft and Wizardry. Children in the UK and Ireland go through generational waves, where each cohort "rediscovers" Harry Potter. Why are some things successful in spreading widely and stably, such as the fictional world of Harry Potter in the last twenty years, while others are not?

An intuitive distinction concerns the effect of social influence versus the features of the content of the traits. The "rediscovery" of Harry Potter is due to parents, elder siblings, and early adopter-peers, from which children learn, as a minimum, about its existence. At the same time, the content of Harry Potter's stories should be attractive enough to reinforce social influence, in order to be stable through years, and in many different countries. Content and social influence are likely to act together, to a different degree. For some traits, however, content seems more important: western children, on average, prefer pizza to boiled spinach; the great majority of cultures use, on some occasions, masks or make-up for faces (1). For others, it may be the opposite: hugging or kissing can be used as greetings in some societies, but considered inappropriate in others where handshakes, or bowing, are used; beanie hats and skinny jeans come and go.

This intuitive distinction reflects important practical differences. Social factors can be leveraged to promote behavioral change (whether for the bad or for the good), while preferences for content features tend to be more stable, as parents or educators trying to have children eating spinach instead of pizza know well. Social factors intuitively should result in more cross-cultural diversity, where relatively unconnected sub-populations converge on different cultural configurations (see e.g. the experimental work in (2)) while more stable content preferences should attract the same sub-populations towards similar outcomes.

The same distinction is used in evolutionary approaches to the study of culture. Epistemic vigilance distinguishes, for example, between the evaluation of the "source" and of the "content" of communicated information (3). In the cultural evolution framework, different mechanisms have been proposed as reflecting social influence, usually under the general label of indirect-biased transmission (4) or context-biased social learning strategies (5): these mechanisms act by selecting among different cultural traits the ones that are associated to some features of the context, or of the population. On the other side, direct-biased transmission, or content-biased social learning, indicates the selection of traits based on their intrinsic features. However, preferences for content also act outside the selection process: we can adopt cultural traits via individual learning, a process sometimes labeled guided variation (4). More generally, various individual processes can make us converge with higher probability to some traits, or particular configurations of traits: in cultural evolution terminology they are referred to as convergent transformation (6, 7), or content-based attraction (8).

In what follows, we present analytical models of cultural evolution that consider both preferences for content and social influence, both with binary and continuous traits (in Supplementary Materials we also describe individual-based versions of the same models). We are using the label "content preferences" for mechanisms that are *not socially influenced*, that is, they do not depend directly on features of the larger population; *directional*, that is, points towards a particular trait, or traits configuration; and, finally, *stable*: they do not change, at least at the time scale of the models. The content preferences modeled here are equivalent, in the binary case, to biased mutation as in (9); in the continuous case, to guided variation (4), biased or convergent transformation (6, 7), or content-based attraction (8).

In opposition, in the models below, "social influence" identifies mechanisms that select traits to copy based on features of the population/source, that are not associated with particular traits, and thus change accordingly to changes in the population. The models consider two of the most studied mechanisms of social influence: conformity (or frequency-based indirect bias) and demonstrator-based indirect bias. Conformity is defined as a disproportionate tendency to copy from the majority, and it is implemented, in the binary case, following (9) for the individual-based version, and (4) for the analytical model. Conformity with continuous traits is rarely modeled: here we follow (10). Demonstrator-based indirect biases instead do not depend on the frequency of traits, but on features of the demonstrator. A classic example is prestige bias, or a tendency to copy preferentially from individuals that are considered to be "high-status" (11), but any tendency that makes preferentially choose some demonstrators because of features independent from the copied traits would fit the description, such as copying preferentially younger (or older) demonstrators. The individual-based models implementing demonstrator-based indirect biases are inspired, with modification, by (9, 12), while for the analytical treatment we follow (13, 14).

Surprisingly, there is not a clear understanding of what are exactly the consequences of dynamics driven by preferences for content or by social influence. Many studies in cultural evolution focus on social influence, possibly because heuristics like conformity or prestige bias can produce population-level adaptive behaviors that go beyond individual cognition, a process that is considered central in cultural evolution (15). Other researchers have instead highlighted the importance of weak but stable preferences for content, as the main way to support cultural transmission and hence stabilize traditions (16). Few studies have considered explicitly the difference between the two processes. (17) showed that when both social influence and preferences for content act, social influence fully determines the outcome. However, their model assumes two preferences for content, and social influence acting stably towards one of the two. In response to this work, (18) presented a model where social influence and the target for content preferences are separated, and they show how the final equilibrium point depends on the relative strength of social influence and content preferences (see also (8)).

In these models, the target of social influence is however fixed and linked to a particular trait, or trait configuration. "Pure" social influence should be instead considered as determined *only* by the context, be it the frequency of any trait in the population, or some demonstrators' features independent from the copied trait. Below, we show that, when this is the case, content preferences, even when weak, determine cultural dynamics, as they are the only directional forces.

Models and results

Frequency-based social influence

The analytical model assumes a large population of individuals. In the “binary trait” model, individuals possess a cultural trait A or B, and p denotes the frequency of A. When examining frequency-based social influence (e.g., conformity), following (4), at each time step, three demonstrators are randomly chosen for each individual. If all have the same trait (three As or three Bs), the individual copies it automatically. In the other cases, the majority trait (i.e., the one possessed by two demonstrators) is adopted with a probability equal to 2/3 + D/3. The parameter D goes between 0 and 1, regulating the strength of conformity. With D=0 (no conformity) the probability of copying the majority trait is 2/3, equivalent to unbiased copying, and with D=1 (maximum conformity) individuals always copy the majority trait. In the models, D is fixed to 1. In addition, individuals have a preference for content: at each time step, with a probability α, individuals switch to trait A.

In our analytical model, we assume that the population is infinite, so that the dynamics are deterministic. The results align with our individual-based simulations using finite populations (see Supplementary Materials). The dynamics of the system can be summarized by the same equation as the one governing conformist transmission in (4), adding a term representing the preference for content:

$$p_{t+1} = p_t + Dp_t(1 - p_t)(2p_t - 1) + \alpha(1 - p_t)$$

When only conformity acts (α=0), the trait that is initially in majority will fixate. With α>0, virtually all equilibria are with populations at the trait A, favored by the content bias. Figure 1 shows the dynamics for all the range of *p*, and for low values of α. When α> ⅛ (0.125), A always fixates. When α<⅛ , we have two stable equilibria: either the fixation of A, or a strong majority of B (> ¾). The one reached depends on the initial value of *p*, above or below the 'unstable equilibria' line in Figure 1, which separates the two basins of attraction.

In the "continuous trait" models, every individual has a continuous trait *p*, bounded between 0 and 1. Following (10), individuals adopts the mean trait in the population, with an error drawn from a uniform distribution between $-\omega(p)$ and $+\omega(p)$, $\omega(p)$ being the standard deviation of the trait in the population. As before, agents have a content preference: with probability α, individuals adopt *p*=1. Again, we assume an infinite population for simplicity, and our results align with our finite population simulations. In this case, the mean *p* of the trait in the population is on average unaffected by the copying process, so that, with α>0, the only possible equilibrium is *p*=1, i.e., for all conditions, given some preference for the content, the population converges on it.

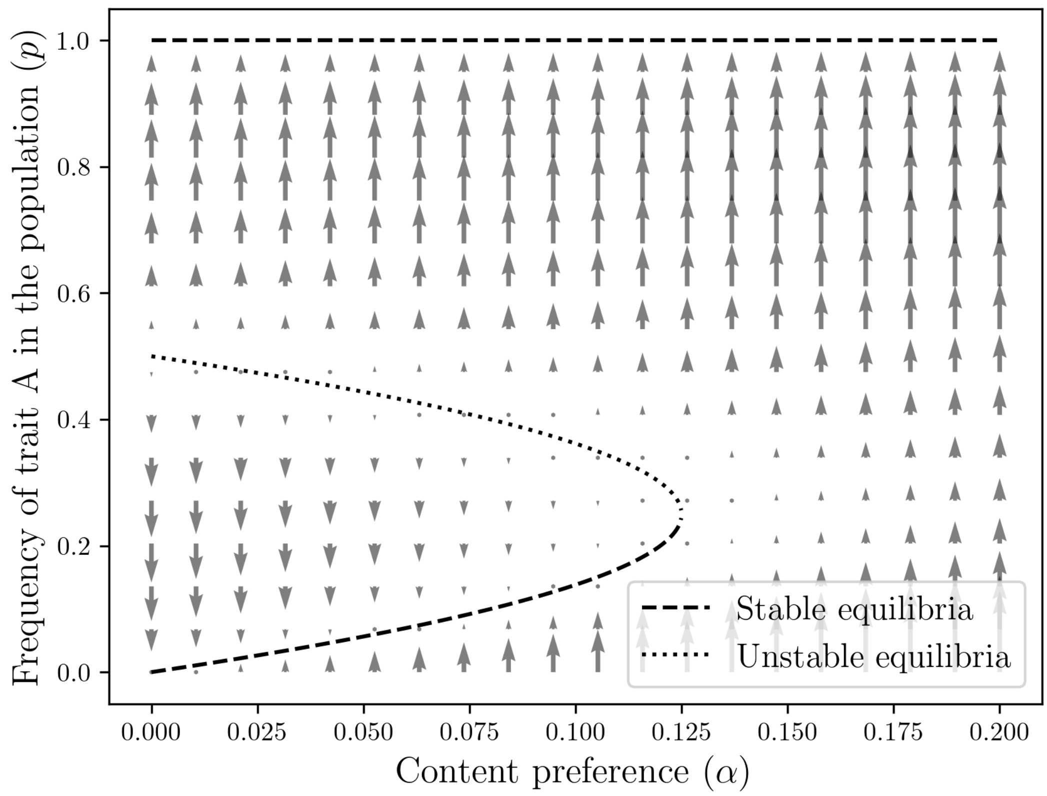

Figure 1. Vector field for the frequency-based social influence model with continuous traits. The trait A always fixate, unless *p* is initially below the unstable equilibrium line and α<1/8 (0.125). For α=1/8, we observe a saddle-node bifurcation: the stable and unstable equilibria collide and disappear.

Demonstrator-based social influence

As above, we consider a population where individuals possess a cultural trait A or B, and *p* denotes the frequency of A. Now, individuals choose only one demonstrator when they update their trait. We assume that a fraction C_s of the population is preferentially chosen (e.g., prestigious individuals), and has a probability C_{copy} times higher to be selected as demonstrators. As before, agents have a preference for trait A: with probability α, they switch to the trait A.

This setting is more subject to randomness: with α = 0, the dynamics resemble a random walk. For this reason, we do not assume an infinite population here, and build instead a stochastic model. When the population is large, the system can be approximated by a Wright-Fisher diffusion (14, 19). There, both the time *t* and the proportion *p* of individuals possessing the trait A become continuous variables, which allows us to make use of differential equations. Then, we can study the long-term behavior of the system, and the time to reach equilibrium (see Supplementary Materials).

The system is governed by the stochastic differential equation:

$$dp_t = \alpha(1 - p_t)dt + \gamma\sqrt{p_t(1 - p_t)}dB_t$$

where B_t is the standard Brownian motion and γ measures the strength of the demonstrator-based bias, encompassing both C_s and C_{copy} (see Supplementary Materials). In other words, the system is subject to two forces, the first one being directional, proportional to the content preference α, and the other one being non-directional, analogous to genetic drift, proportional to the demonstrator-based bias strength γ.

As shown in (20), In the long run, the trait A always fixates: *p*=1 is the only absorbing state of the equation. A possibly more surprising result is that the stronger the demonstrator-based bias (either because individuals that are preferentially copied are rarer, or because they are more influential), the quicker the fixation (see Figure 2). In our model, individuals that are preferentially copied do not have, on average, a different trait than the rest of the population. As visible in the above equation, the demonstrator-based bias has therefore no directional effect, but only increases the volatility by shrinking the pool of demonstrators, which makes it easier for the favored trait to fixate.

If the trait is instead continuous, we can apply the same reasoning as in the frequency-based social influence case (see Supplementary Materials). As long as α > 0, the trait A fixates.

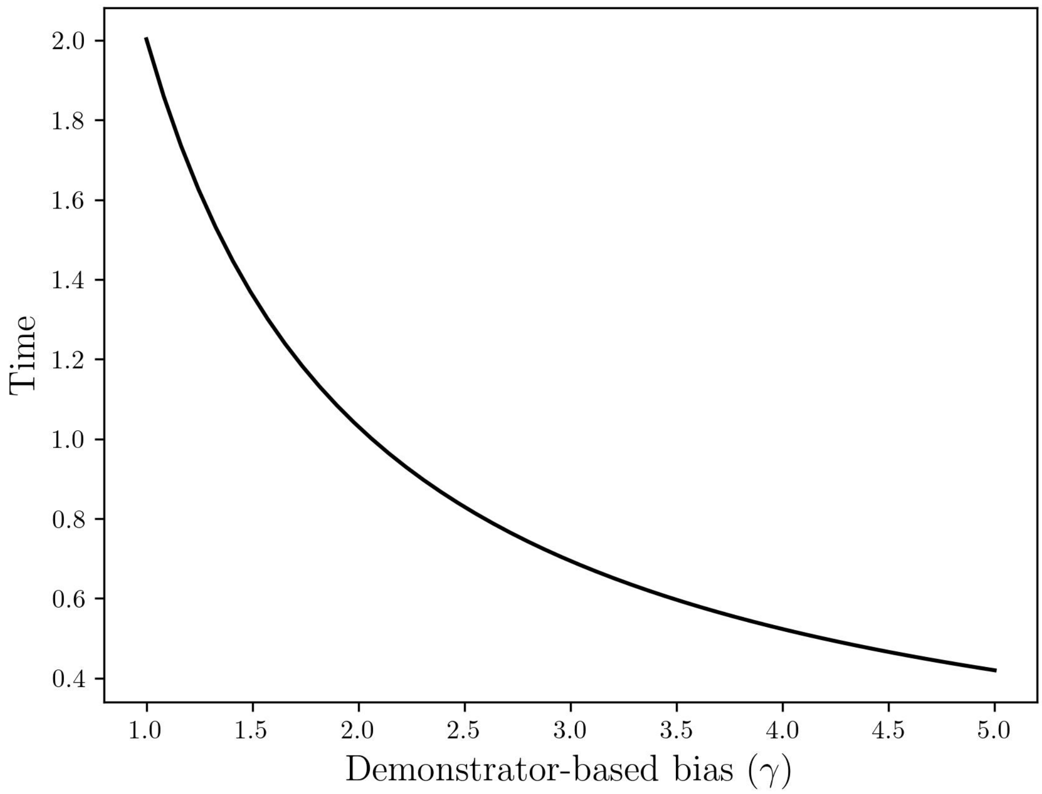

Figure 2. Time to equilibrium in function of the strength of demonstrator-based social influence. The stronger the demonstrator-based bias (γ), the faster the fixation of the trait favored by the content preference (here α=.1 and p is initially .5).

Discussion

The results of the models show that, in virtually all scenarios, content preferences stabilize culture at the point where they are directed to. Social influence mechanisms - both based on frequency (conformity) and on demonstrators' features (e.g., prestige bias) -, are independent from traits' features, and non-directional. When they act together with content preferences, the latter are therefore the only directional forces, and their existence is sufficient to determine the fate of the system.

This message suffers one – interesting – exception. When a majority of the population holds the non-favoured trait in a binary choice, and conformity acts, the preference for the (minority) traits needs to be sufficiently strong to overcome the majority. In fact, conformity is non-directional on average, but reinforces existing majorities. In cultural evolution, it has long been recognised that a conformist bias can make cultural traits persistent, and 'maintain between-group cultural variation' (20, 21). But if a content preference for trait A builds gradually, our model suggests that cultural change could happen suddenly: the trait A would stay rare for a while, then suddenly spread as it crosses 1/8 (Figure 1). When conformity and content preference act together, a subtle change in preferences can be enough to trigger a sudden cultural shift. Our model thus provides a possible parsimonious explanation for 'tipping points' in cultural evolution (23, 24).

As mentioned in the Introduction, few works have explicitly addressed these questions, but the results presented here are consistent with suggestions coming from other cultural evolution models, and give them a more general background. (6), for example, found that convergent transformation drives cultural dynamics when acting together with unbiased copying (and, similarly to here, the more faithful the copying is, the stronger the effect of convergent transformation). (10) found that even weak priors render conformity unable to stabilize traditions and determine the outcome, in most conditions for binary choices, and always for continuous choices (these results are echoed in the more recent (25)).

The immediate take-home message of these results is that, if our question is why some things are culturally successful and others are not, weak but stable non-social forces need to be taken into account. A possibly less obvious take-home message concerns the interpretation of cross-cultural regularities. The existence of human universals (26) is sometimes interpreted as supporting the existence of strong cognitive evolved dispositions, or strong ecological constraints, and indicating a somehow limited role of culture. On the other end of the spectrum, socio-cultural anthropologists have tended to diminish the importance of cross-cultural regularities to stress the importance of culture. While everyone would agree this is a false dichotomy (see e.g., (27)), these results suggest a way to understand why it is so: weak directional, non-social forces, as long as they are stable enough, can produce strong regularities. These can be (possibly weak) cognitive priors, physical affordances, relatively stable ecological conditions (such as the availability of certain materials), and so on. Conceptually, it is important to think of social influence and preferences for content not as opposing forces. Non-directional social influence provides strength to the weak but directional preferences for content. In other words, culture magnifies individual-level tendencies, allowing them to become stable at population level. This can be clearly seen, in our model, in the case of the demonstrator-based bias, where the stronger is the social influence, the faster is the convergence towards the equilibrium to which content preferences point.

In the models presented here, the content preference is uniform in the population, i.e., only one preference was considered, but the same logic applies to more realistic situations with many different forces, and we would expect culture to homogenize population towards the stronger ones. In addition, the preference for the content is, in the models, deterministic: individuals that are subjected to it (according to α) necessarily switch to the preferred trait. Again, this is a simplification, but a probabilistic implementation would not change qualitatively the results, only possibly producing longer times to reach fixation at the preferred trait.

Finally, the modeled social influence mechanisms are fully detached from the content of the traits. In reality, we expect that, for example, prestigious individuals would possess, on average, more adaptive traits than individuals chosen at random, or that the majority targeted by conformist copying would effectively pool information from individual learning (4). However, we believe this reinforces our point, as underscore again how social influence needs to be guided by the features of traits to be effective.

Acknowledgements:
James Winters suggested to implement copying probabilities in the demonstrator-biased simulation with a softmax function. Marius Mercier helped spotting typos in the preprint. José Moran and Emilien Jemelen helped with the analytical model.

References

1. D. Sperber, L. A. Hirschfeld, The cognitive foundations of cultural stability and diversity. *Trends Cogn. Sci. (Regul. Ed.)* **8**, 40–46 (2004).
2. M. J. Salganik, P. S. Dodds, D. J. Watts, Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science* **311**, 854–856 (2006).
3. D. Sperber, *et al.*, Epistemic vigilance. *Mind & Language* **25**, 359–393 (2010).
4. R. Boyd, P. J. Richerson, *Culture and the evolutionary process* (University of Chicago Press, 1985).
5. R. L. Kendal, *et al.*, Social learning strategies: bridge-building between fields. *Trends in cognitive sciences* **22**, 651–665 (2018).
6. A. Acerbi, M. Charbonneau, H. Miton, T. Scott-Phillips, Culture without copying or selection. *Evolutionary Human Sciences* **3**, e50 (2021).
7. A. Mesoudi, Cultural selection and biased transformation: two dynamics of cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* **376**, 20200053 (2021).
8. N. Claidière, *et al.*, Convergent transformation and selection in cultural evolution. *Evolution and Human Behavior* **39**, 191–202 (2018).
9. A. Acerbi, A. Mesoudi, M. Smolla, *Individual-Based Models of Cultural Evolution: A Step-by-Step Guide Using R* (Routledge, 2022) https:/doi.org/10.4324/9781003282068.
10. T. J. H. Morgan, B. Thompson, Biased transformation erases traditions sustained by conformist transmission. *Biology Letters* **16**, 20200660 (2020).
11. Á. V. Jiménez, A. Mesoudi, Prestige-biased social learning: current evidence and outstanding questions. *Palgrave Communications* **5**, 1–12 (2019).
12. A. Mesoudi, The Cultural Dynamics of Copycat Suicide. *PLOS ONE* **4**, e7252 (2009).
13. J. Moran, A. Fosset, M. Benzaquen, J.-P. Bouchaud, Schrödinger's ants: a continuous description of Kirman's recruitment model. *Journal of Physics: Complexity* **1**, 035002 (2020).
14. A. Etheridge, *Some Mathematical Models from Population Genetics: École d'Été de Probabilités de Saint-Flour XXXIX-2009* (Springer, 2011).
15. J. Henrich, The Secret of Our Success (2016).
16. O. Morin, *How Traditions Live and Die* (Oxford University Press, 2015).
17. J. Henrich, R. Boyd, On Modeling Cognition and Culture: Why cultural evolution does not require replication of representations. *Journal of Cognition and Culture* **2**, 87–112 (2002).
18. N. Claidière, D. Sperber, The role of attraction in cultural evolution. *Journal of Cognition and Culture* **7**, 89–111 (2007).
19. W. Feller, Diffusion processes in genetics in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, (University of California Press, 1951), pp. 227–247.
20. A. Mesoudi, Migration, acculturation, and the maintenance of between-group cultural variation. *PLOS ONE* **13**, e0205573 (2018).
21. J. Henrich, R. Boyd, The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. *Evolution and Human Behavior* **19**, 215–241 (1998).
22. J. J. Bissell, Conformity bias and catastrophic social change. *Tipping Points: Modelling Social Problems and Health*, 168–182 (2015).
23. D. Centola, J. Becker, D. Brackbill, A. Baronchelli, Experimental evidence for tipping points in social convention. *Science* **360**, 1116–1119 (2018).
24. O. Kolodny, N. Creanza, M. W. Feldman, Evolution in leaps: The punctuated

accumulation and loss of cultural innovations. *Proceedings of the National Academy of Sciences* **112**, E6762–E6769 (2015).
25. M. Yan, S. Mathew, R. Boyd, “Doing what others do” does not stabilize continuous norms. *PNAS Nexus*, pgad054 (2023).
26. D. E. Brown, *Human Universals* (new York: McGraw Hill, 1991).
27. H. C. Barrett, *The Shape of Thought* (Oxford: Oxford University Press, 2015).