<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *Emotion Review*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Eriksson, K., Strimling, P. (2020)

Using Models to Predict Cultural Evolution From Emotional Selection Mechanisms

Emotion Review, 12(2): 79-92

<https://doi.org/10.1177/1754073919890914>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:mdh:diva-46921>

Using Models to Predict Cultural Evolution from Emotional Selection Mechanisms

Kimmo Eriksson

Stockholm University and Mälardalen University

Pontus Strimling

Stockholm University

Corresponding author:

Kimmo Eriksson

Center for Evolutionary Culture, Stockholm University

Wallenberglaboratoriet

106 91 Stockholm

SWEDEN

Email: kieri@ark.su.se

Abstract

Cultural variants may spread by being more appealing, more memorable, or less offensive than other cultural variants. Empirical studies suggest that such “emotional selection” is a force to be reckoned with in cultural evolution. We present a research paradigm that is suitable for the study of emotional selection. It guides empirical research by directing attention to the circumstances under which emotions influence the likelihood that an individual will influence another individual to acquire a cultural variant. We present a modeling framework to translate such knowledge into specific and testable predictions of population level change. A set of already analyzed basic cases can serve as a toolbox.

Keywords: cultural evolution; emotional selection; imitation; social sanctions; social norms

Using Models to Predict Cultural Evolution from Emotional Selection Mechanisms

Human culture changes over time, whether considered in terms of behavior, values, ideas, technology, etc. This process of change, through the selection of one cultural variant over another, can be regarded as an instance of evolution. As it does not require any change in humans' genetic makeup, it is referred to as *cultural evolution*. An abundance of intriguing cultural evolutionary phenomena can be observed. For instance, hygiene norms have become stricter throughout recorded time; written language has been invented several times, and spread throughout the world to become a necessary part of society; ideas concerning how society should be fair, free, and democratic have become dominant in many countries; and an enormous number of tools and technologies have spread throughout human history. The aim of the research field of cultural evolution is to understand the processes underlying such cultural changes.

In this paper, we think of culture as a distribution of “cultural variants” in a population. A cultural variant is a unit of culture that is carried by an individual, such as an idea, a habit, a tool, etc. Cultural evolution involves the creation of novel cultural variants that may spread in the population at the expense of other cultural variants. Typically, a cultural variant needs to be expressed by the individual in order to be observable to others. If other people tend to pick up the cultural variant and favor its spread, it will become more common in the population. This account of cultural evolution has been termed the “epidemiological approach” (Nichols, 2002; Sperber, 1996). In essence, the epidemiological approach asks researchers to figure out what aspects of certain cultural variants make them successful in cultural evolution. The answer to this question may often lie in the emotional impact of the cultural variants. In this paper, we take the epidemiological approach and focus on the role of emotions.

Emotional selection

To illustrate the role of emotions in the epidemiological approach, we shall consider a couple of classic examples from the literature. Our first example is the class of cultural variants known as urban legends, that is, generic anecdotes about someone to whom something happens. People who hear an urban legend may then retell it to others, but often they don't. This process of selective retelling will cause cultural change. Only those urban legends that, for whatever reason, tend to be retold will spread in the population. What might those reasons be? Two alternatives were formulated in a pioneering study of urban legends by Heath, Bell and Sternberg (2001); urban legends might be retold based on the quality of their informational content ("informational selection"), or based on the emotions they create ("emotional selection"). The authors found strong support for emotional selection, specifically with regard to the emotion of disgust. Later work has found that emotional selection of urban legends can be based also on other arousing emotions, such as happiness, surprise, anger, and anxiety (Berger, 2011; Berger & Milkman, 2012; Eriksson, Coultas, & De Barra, 2016; Fessler, Pisor, & Navarrete, 2014; Peters, Kashima, & Clark, 2009). Moreover, emotions affect not only which stories people choose to pass along, but also which stories they choose to hear, and which stories they recall (Eriksson & Coultas, 2014). The effect of emotional arousal on memory has, of course, long been known (Heuer & Reisberg, 1992). The point here is that this mechanism could contribute to cultural evolution. If more emotionally arousing stories are more memorable, they will be more likely to survive and spread in a population.

The spread of urban legends fits the notion of emotional selection well, but it is a rather peripheral phenomenon of cultural evolution. As a more central phenomenon, consider

the evolution of social norms. Nichols (2002) noted that centuries-old books of etiquette contain prohibitive norms that are still valid today (e.g., that you should not urinate in public), as well as prohibitive norms that are not valid any longer (e.g., that when sitting down to eat you should not clasp your hands together). Nichols took the epidemiological approach to understand what aspect of an action would make a prohibitive norm against it more likely to survive. He argued that norms against actions that involve bodily fluids, and hence tend to elicit core disgust, are more easily remembered and felt to be more serious and important than affectively neutral norms. Consistent with this argument, Nichols found that the relation to core disgust predicted which prohibitive norms in an old book of etiquette are still valid today.

Medical practices may also be subject to emotional selection. Miton, Claidière, and Mercier (2015) argued that this is the case for bloodletting, a treatment that does not help the patient, but nonetheless has been practiced in many cultures across the world. It may simply feel cleansing when “bad” blood pours out, giving bloodletting an advantage in cultural evolution. Another mechanism whereby ineffective treatments may spread through emotional selection is discussed by De Barra, Eriksson, and Strimling (2014); when a treatment has no systematic effect, there is still a random variation in outcomes. An individual who has experienced a positive outcome may be more motivated to talk about it. This bias, in which outcomes of a treatment become known, can give ineffective treatments a better reputation than they deserve.

This review of empirical case studies on emotional selection is not exhaustive. Still, it should be sufficient to warrant the conclusion that emotions may play a crucial role in the mechanisms underlying cultural change. See also Elster’s (2015) book on how to explain social behavior, which devotes an entire chapter to the key role of emotions. We now turn our

attention to formal models that can predict cultural change, and the question of how such models can incorporate emotion-based mechanisms.

The purpose of formal models of cultural evolution

Inspired by the success of mathematical models in population genetics, the use of mathematical models of cultural evolution became popular in the 1980s (Boyd & Richerson, 1988; Cavalli-Sforza & Feldman, 1981; Lumsden & Wilson, 1981). This body of research is mainly focused on how biological and cultural evolution interact to create behavioral and genetic change over evolutionary time. Here, we are concerned instead with models of cultural evolution as a tool to help us explain specific cases of cultural change in historic time. The purpose is thus to contribute to the study of questions that are within the domain of mainstream social science. For instance, social norms and how they change constitute a central theme in sociology. As mentioned above, Nichols (2002) used the epidemiological approach to contribute to our understanding of these changes. However, Nichols did not use any mathematics to present this evolutionary theory. In fact, the use of mathematics in social science may sometimes serve no higher purpose than to make a research paper look more rigorous (Eriksson, 2012). Nonetheless, we shall argue that formal models of cultural evolution may be genuinely useful.

Our conception of culture as a distribution of cultural variants in a population is inherently quantitative. In this conception, phenomena of cultural change can be described in mathematical terms as the increase or decrease, at certain rates, of the frequency of certain cultural variants. Therefore, it makes sense to also try to describe the causal process in mathematical terms.

A mathematical model has some advantages over a purely verbal theory, such as its being more transparent about its underlying assumptions. A mathematical formulation also makes it easier to examine how predictions from the theory rely on various assumptions. Moreover, a mathematical model may yield predictions with a higher precision than a verbal theory can produce. These advantages of formalizing a theory into a mathematical model have little value in themselves, however. A formal model is just a representation of a theory—if the theory lacks validity, so does the model. On the other hand, if a theoretical idea has validity, then theorists might increase its value further by formalizing it. We believe that the notion of emotional selection is valid, and may therefore benefit from formalization.

Modeling emotional selection

So how can we make formal models of dynamics based on emotional selection? As a starting point, we make a very brief inventory of two popular paradigms for modeling cultural evolution. One influential paradigm is based on game theory, which means that cultural variants are assumed to yield some kind of payoff to the individual and to others, and that the distribution of cultural variants changes as a consequence of some optimization of this payoff (see Cownden, Eriksson, & Strimling, 2017). A related paradigm, known as gene-culture coevolution, incorporates various transmission biases, but the focus is still on how cultural variants pay off in terms of functionality or fitness (Boyd & Richerson, 2005; Richerson & Boyd, 2005).

These modeling paradigms are not well-suited for capturing the role of emotions in cultural evolution. As Heath et al. (2001) pointed out, a cultural variant, like an urban legend, may not have any clear functionality at all, and even if it has, functionality is not necessarily central to the mechanism by which it spreads. Nichols' (2002) postulation that prohibitive

norms regarding certain actions, regardless of whether they optimize any payoff, may spread simply because emotions about the actions make the norms easy to remember, and make them *feel* important, is one example of this. The role that emotions play could well be evolutionarily shaped, so that people on average have benefited from, say, taking particular notice of disgusting things. The point here is that, to understand why a specific cultural variant spreads, it may be the emotional response it elicits that we need to consider, rather than any benefits specific to that cultural variant.

To capture emotional selection, we shall instead turn to another paradigm, one that does not depend on payoffs to understand change. We shall refer to the level of individuals as the *micro*-level, and the level of the population as the *macro*-level (Schelling, 2006). We shall conceive of the process of cultural evolution as a series of micro-events, in which individuals change their cultural variants, which may happen many times over an individual's lifetime (Strimling, Enquist, & Eriksson, 2009). Specifically, we shall assume that these micro-changes are mainly due to various kinds of social influence, so that they can be attributed to *interactions* between agents. The outcome of an interaction is that one or more of the parties may change a cultural variant. For example, an interaction may involve one party retelling an urban legend that may then be added to the other party's repertoire of anecdotes. Of course, we cannot have detailed knowledge of all the micro-events that take place in a human population. When knowledge is lacking, the mathematical solution is to assign *probabilities* of various micro-events occurring. This yields a very general modeling paradigm for cultural change that is well-suited to dealing with emotional selection. The effect of emotions is then represented in the probability that a cultural variant is shown, acquired, or remembered.

Representing the effect of emotions as probabilities

In sum, we are advocating for a modeling paradigm that requires researchers to figure out the circumstances under which some kind of social influence may cause individuals to change their cultural variant in this domain, and to represent this knowledge as probabilities of various change events.

We shall now discuss how this could be done, based on a conceptualization of social influence as involving two roles: someone is potentially influencing someone else to change their cultural variant. For a more easily digestible terminology, we shall use “send” and “receive” instead of “potentially influence” and “potentially be influenced,” and refer to the roles as sender and receiver. Note that no intention is required to be a sender of a cultural variant; for instance, simply by wearing a piece of clothing, you may make it visible to, and thereby potentially influence, a receiver. In a one-sided interaction, only one party can be influenced. In a two-sided interaction, both parties may influence each other. Thus, in a two-sided interaction, an individual may simultaneously act as sender and receiver.

To illustrate the kind of work that is required to represent the effect of emotions as probabilities, consider imitation. How could emotions play a role in the probability of change due to imitation? For one thing, the sender must display the behavior for imitation to occur. For another, the receiver must adopt the behavior on display. The probability of a successful change event is the probability that both of these things happen. Thus, imitation will be more likely if the behavior is one that people like to do a lot when observed, and that observers easily remember and feel like trying themselves. What researchers need to do is examine how emotions affect these things. An example we have discussed earlier is how arousing content may make an anecdote more likely to be told, more likely to be listened to, more easily remembered, and more likely to be retold. Other emotions may, instead, work as inhibitors of

imitation. For instance, behaviors that people tend to think are shameful will be less likely to be displayed in public, hence less easily imitated.

Of course, imitation is just one of many possible mechanisms for social influence. In the case studies discussed below, we will present two other mechanisms whereby emotions might affect change: argumentation and social sanctions. We thus need to consider people's emotions towards different arguments, and the emotions that may trigger social sanctions.

The value of taking change events into account

Temporarily setting emotions aside, we also want to emphasize the general importance of explicitly taking change events into account - even when selection is based on optimizing payoffs. This should be of interest to those readers who are already acquainted with some evolutionary modeling. To illustrate the importance of explicitly describing change events, we shall show how the outcome may depend on exactly how individuals use change events to optimize payoffs.

Consider the famous "stag hunt" game, a situation studied in game theory as an example of a choice between a risky strategy and a risk-free strategy (Skyrms, 2004). Two individuals ("players") from a larger population are paired up and make separate decisions on whether to hunt stag or hare. Hunting stag is assumed to be a risky strategy, in that it has a high payoff if both players choose to do, it but no payoff if one does it alone. Hunting hare is assumed to be a risk-free strategy, in that it gives a medium payoff regardless of what the other player does. How will the proportions of the two strategies evolve in a population?

A standard approach to this question, among evolutionary game theorists, is to assume that strategies spread among the population in relation to their expected payoff. This means that individual change events are not explicitly considered; researchers take a leap from

payoffs to an assumption about population-level change. Under this assumption, mathematical analysis shows that evolution will be path dependent: if the initial proportion of stag hunters is large enough, the population will evolve toward only hunting stag, otherwise it will evolve toward only hunting hare (e.g., Sandholm, 2009). However, a different conclusion may be drawn if we instead take the specifics of change events into account. The outcome then depends on how people go about changing their strategies to optimize their payoff, as illustrated by the following two simple cases.

Case 1: Play first, then compare payoffs and update your strategy. Suppose an interaction involves the two players first playing according to their current strategies, and then comparing the payoff. If the payoffs are different, the player who received the lower payoff will use a different strategy in the next interaction. As the stag hunt is symmetric, the two players can only get different payoffs when one player hunts stag and the other hunts hare, in which case the stag hunter gets the lower payoff and changes strategy to hunting hare next time. The dynamic consequence is that the strategy of hunting hare will increase in frequency until *everyone is hunting hare*.

Case 2: First compare payoffs and update your strategy, then play. Now suppose an interaction instead involves the two players first comparing the payoffs they received in the previous round and, if different, the player who received the lower payoff changes strategy when the two players proceed to playing the game. Any player who has hunted stag together with another player will have received the maximal payoff. In these players' next interaction, any players they interact with who have previously hunted hare will always change to hunting stag instead, and both will receive the maximal payoff. Thereby, the set of players hunting stag increases. The dynamic consequence is that the strategy of hunting stag will always increase until *everyone is hunting stag*.

Thus, even in a domain where population change is only driven by individuals changing to maximize their payoff, it is crucial to consider the exact circumstances under which they do this.

Connecting models with empirical research

The micro-to-macro modeling paradigm that we advocate should connect with two levels of empirical research in a given domain. First, at the micro-level, the model needs to connect with research on the mechanisms whereby changes of individuals' cultural variants tend to occur in this domain, and how these mechanisms are sensitive to differences between cultural variants and between individuals. For instance, suppose researchers conclude that social sanctions is the main mechanism for change in a given domain. The left column of Table 1 presents a number of empirical micro-level questions regarding social sanctions in the given domain that need to be considered. These are questions that we hope will interest emotion researchers.

Table 1. Examples of empirical questions asked at different levels.

Micro-level	Macro-level
Which cultural variants in the given domain are most likely to evoke emotions that motivate people to use social sanctions?	How does culture in the given domain vary over time?
Which individuals are most likely to use social sanctions?	How does culture in the given domain vary across (sub)populations?
How do people emotionally respond to social sanctions?	How does cultural change vary across different cultural variants in the same domain?
Which individuals are thereby most likely to change their cultural variants due to social sanctions?	

A modeler can then set up a mathematical model with probabilities informed by the answers to the previous questions. The bulk of this paper is devoted to describing how a model can be set up and analyzed. The results of this analysis are predictions from the model regarding how the frequency of different cultural variants should change in the population over time, and how this should vary depending on the specific features of cultural variants and individuals.

Finally, at the macro-level, these predictions need to connect with empirical research on the cultural phenomenon of interest. This requires answers to macro-level questions of the kind given in the right column of Table 1. To the extent that model predictions are consistent with macro-level data, the researchers will have achieved an account of a cultural phenomenon that is grounded in validated micro-level mechanisms.

Jon Elster is a leading philosopher of social science. The approach we describe here is close to Elster's (2015) ideal of social science, in which "triggering probabilities" give us enough insight to specify a distribution of behavior in a population, and thereby predict the aggregate behavior. However, Elster also warns that this ideal may not be achievable because of the importance of a complex web of contextual details in determining the probabilities for change (Elster, 2015, p. 466-467). To demonstrate that this ideal may, in fact, sometimes be achieved, we will briefly discuss two recent case studies published in *Nature Human Behaviour*: a study of change in hygiene and violence norms (Strimling, De Barra, & Eriksson, 2018); and a study of change in moral opinions (Strimling, Vartanova, Jansson, & Eriksson, 2019).

Lastly, we turn to a description and discussion of the mathematical modeling framework. This section includes a catalog of basic cases that we have already analyzed. If

empirical researchers identify that micro-events in some domain maps to one of these basic cases, they can use this catalog to make predictions about the macro-level. This too is illustrated by the case studies below, which mapped to two different basic cases.

Case Studies

Hygiene and violence norms

An intriguing cultural phenomenon is that the use of violence, and the handling of bodily fluids, have tended to become more restricted over time. These trends were first documented by Norbert Elias, who called them the “civilizing process” (Elias, 1982; Elias & Jephcott, 1978; Pinker, 2011). Elias and others have suggested various drivers of the civilizing process. They typically involve sociological features specific to certain times and places. As the phenomenon seems to prevail in a much greater generality, a more generally applicable explanation seems to be required. The theory of Nichols (2002), that disgust affects which rules we remember, is not sufficient to explain the phenomenon, as it pertains to what rules remain rather than how more restrictive rules are able to spread. The theory is also incomplete from the point of view of our modeling paradigm. It does not address what alternative cultural variants that individuals can hold for a specific norm, nor the circumstances under which individuals change.

In a recent paper, we proposed that norms address behavior for which there are typically two different views present in the population: some individuals think the behavior should be prohibited; whereas other individuals think that it should *not* be prohibited (Strimling et al., 2018). In other words, people have different preferences for what the norm should be. Some people endorse a stricter norm, others endorse a less strict norm. We further proposed that change in the domain of norms is likely to be driven mainly by social

sanctions. When two individuals with different norms interact, both may regard the other as a norm violator who possibly should be punished. Emotional selection will then occur if differing norm violations elicit negative emotions of differing strengths, so that punishment is more likely to go in one direction than in the other.

On the micro-level, we required empirical evidence for asymmetry in the elicitation of emotions and use of punishment. Such evidence was collected through surveys on behaviors for which there is substantial norm heterogeneity in the population, such as hand washing before a meal, and whether to sneeze into a tissue. The survey data supported the hypothesis that behavior that deviates from one's own norm in the direction of being less hygienic or more violent tends to create stronger negative emotions (feelings of threat) than does behavior that deviates from one's own norm in the other direction. Furthermore, it showed that this emotional asymmetry lead to a punishment asymmetry, meaning that less hygienic and more violent behaviors tend to be targeted with more social sanctions. Based on these micro-level mechanisms, our model of the cultural evolutionary process predicts the macro-level phenomenon that norms regarding hygiene and violence should, in general, tend to shift toward becoming increasingly restrictive. This is the macro-level phenomenon that Elias and others have already documented.

Moral opinions

In the domain of public opinion, the cultural variants are the different opinions one can hold on various issues. An empirically established phenomenon of cultural evolution in this domain is that, for as long as polling of moral opinions has existed in the United States, public opinion on moral issues has become overall more liberal. Emotional reactions seem to play a major role in how people judge what is morally wrong or right (Haidt, 2007). In a

couple of recent papers, we used this insight in an attempt to understand why moral opinions change over time (Eriksson & Strimling, 2015; Strimling et al., 2019). The trend toward more liberal opinions suggested that we should look at what differentiates liberal opinions from conservative opinions. A core finding of researchers on “moral foundations” is that conservatives and liberals differ in their emotional responses to different moral arguments. Liberals feel that harm and fairness are singularly important considerations, whereas conservatives feel that it is equally important to consider authority, loyalty, and purity (Graham, Haidt, & Nosek, 2009). From this core finding, it seemed a natural next step to examine each opinion on any given moral issue in terms of which moral arguments can be used to support it. On any issue, we expected the liberal opinion to be the one that was best supported by harm and fairness arguments. A survey demonstrated that this was indeed the case.

Based on these findings, we proposed that change in the domain of moral opinions is likely to be driven mainly by moral arguments. We modeled how the frequency of opinions in a population may change through an exchange of moral arguments between individuals, under the assumption that a typical liberal should be easier to sway to the position that is supported by arguments related to harm and fairness (whereas a typical conservative, feeling that all kinds of moral arguments carry the same weight, should be roughly equally easy to sway to either position). The model predicts that the cultural evolution of public opinion should favor those moral positions supported by arguments related to harm and fairness. These issue positions should tend to become more popular over time, both among liberals and conservatives, but with liberals leading the trend. Evolution should go fastest on those issues for which there is the clearest difference between different positions in terms of how well they are supported by harm-and-fairness arguments.

Note that this model connects with two kinds of micro-level data, on individuals and on moral positions. The data on individuals demonstrates that liberals and conservatives feel differently about moral arguments (Graham et al., 2009). Our survey data on moral positions measures how different positions on the same issue differ in regard to the moral arguments that support them. On the macro-level, we can likewise connect the model to very rich polling data (from the General Social Survey) on how opinions on a large set of moral issues have changed among liberals and conservatives during the last 40 years. The model describes a process for how the micro-level mechanism should drive the macro-level phenomenon of public opinion change. As the model predicts, both liberals and conservatives have gradually shifted their opinions in the liberal direction over time, and more rapidly for issues with a greater difference between different positions in terms of how well they are supported by harm-and-fairness arguments. See Figure 1.

Figure 1. Opinion change toward more liberal opinions, among liberals and conservatives alike, has been faster for issues where positions greatly differ in their supporting arguments. The graphs show, for two halves of a set of 74 moral issues, the time

trends of the average proportions of liberals and conservatives that hold the more liberal opinion. The left (vs. right) panel shows the trends on moral issues where opposing positions exhibit a small (vs. big) difference in how well they are supported by harm-and-fairness arguments (HF). For details, see Strimling et al. (2019).

Conclusion from case studies

The two case studies above showcase a research paradigm. Its goal is to understand cultural evolutionary phenomena from the bottom up. To achieve this goal, researchers are required to examine the micro-level mechanisms of cultural change, including the effects of emotions on the likelihood that particular cultural variants are sent and received. The second step, based on this understanding, is to conduct a theoretical analysis of the process whereby these mechanisms lead to change at the population level. The last step is to compare the predicted outcomes of this process with the actual population-level data.

Note that researchers are not required to examine every single micro-level mechanism by which cultural variants can change. To explain cultural change, we only need to consider those mechanisms that have different effects on the different cultural variants, that is, those having an asymmetric influence on the cultural change by giving more benefit to one of the traits. Any mechanism that has the same effect on all cultural variants, on average, may be disregarded. In other words, the challenge is to pinpoint the asymmetries at the micro-level that drive population-level change. In the first case study, we identified that different acts in the domains of hygiene and violence exhibited an asymmetry in how likely they are to elicit strong negative reactions: too unhygienic or too violent behaviors tend to elicit stronger negative reactions than do behaviors that are too hygienic or too nonviolent. In the second case study, we identified that different opinions on a moral issue may exhibit an asymmetry

in how effective their supporting arguments are: arguments based on harm and fairness are more generally effective than arguments based on authority, loyalty, and purity. The fact that we only need to consider the mechanisms that have an *asymmetric* influence on cultural change is the reason why we can disregard most context, thereby overcoming Elster's problem that we discussed earlier. For instance, suppose that people are more likely to pick up cultural variants from their friends than from strangers. This mechanism may nonetheless be disregarded when modeling cultural change, unless friends and strangers also differ systematically in their cultural variants.

In addition to contributing to cultural evolution research, this paradigm may create novel research questions and answers, both at the micro-level and the macro-level. For example, the moral opinions study required us to consider what it is about a given moral opinion that makes it liberal or not. Our novel answer to this micro-level question was that the key property is the extent to which harm and fairness arguments (which liberals feel most strongly about) support the given opinion rather than the opposite opinion. This answer, when used as an input to our cultural evolutionary model, yielded a prediction that led to a novel empirical insight at the macro-level: the rate of change in public opinion varies between moral issues according to the extent to which harm and fairness arguments support one opinion, rather than the opposite opinion.

A Micro-to-Macro Modeling Framework

The research paradigm discussed above has three steps. The first and the last steps (i.e., the empirical examinations at the micro- and macro-levels) are highly domain-specific. By contrast, the modeling step that derives macro-level predictions from micro-level assumptions is domain-general. The rest of this paper focuses on this modeling step. Our aim

is, first, to explain the concepts of our modeling framework and, second, to provide a catalog of already analyzed basic cases for empirical researchers to refer to. We have also tried to make this section digestible for readers without a mathematics background. For mathematically inclined researchers who want to know the nuts and bolts, and perhaps go beyond these basic cases, we include a mathematical treatment in the supplementary material.

Assumptions

A key assumption is that the cultural trait we are studying is such that every individual “bears” exactly one of several *cultural variants*. Particularly common is the case where we think of a cultural trait as having exactly two variants, which includes all cases when an individual either does or doesn’t behave in a certain way. The term “cultural” underscores that we assume that the trait is not innate; individuals are assumed to choose a cultural variant through social processes, and the same individual can later change to a different variant through the same process.

We are interested in how people react to cultural variants. It may not be the case that everyone reacts in the same way. To the extent that such individual differences are relevant to include in the model, we do so by considering individuals as belonging to different *types*. For simplicity, these types are assumed to be constant. (If the reaction is instead something individuals change through social learning, it would be better to model it as a cultural variant in its own right.) Although we here use types to represent individual differences in emotional reactions, types could also be used to represent other individual differences (e.g., location, power, resources), or different kinds of agents (e.g., individuals vs. organizations) that a researcher thinks is relevant to a given case study. We are interested in how the distribution

of different types in the population influences cultural evolution, in the population as a whole, as well as within each type.

The model focuses on changes of cultural variants that occur as a result of social influence. To simplify, social influence is here assumed to happen in social interactions of pairs of individuals. (For some domains, it may be that the main context in which social influence happens cannot usefully be modeled as pairwise interactions, in which case an alternative modeling assumption should be made.) In a social interaction, individuals may be influenced by each other. We think of someone who is influenced as a *receiver*, and someone who influences as a *sender*. Note that this is a distinction between roles, not between individuals; the same individual may simultaneously hold both roles, that is, influence and be influenced at the same time. Knowledge regarding the circumstances under which individuals may change their cultural variants is represented in the model as *probabilities* of a given social interaction resulting in the receiver changing to the sender's cultural variant. This change will typically depend on the receiver's and sender's types, and their current cultural variants.

We further need to assume that the population is large. The reason for this assumption is that we want to be able to predict macro-phenomena from micro-mechanisms. This means that we need to be able to generalize fairly confidently from individual tendencies (probabilities) to the population level (frequencies). This cannot be done if the population is small, but it can be done if the population is large, a phenomenon known as the "law of large numbers."

Finally, which pairs of individuals tend to interact must be specified. If the modeler has reason to believe that certain types are less likely to interact due to some kind of social

distance, this can be taken into account in the model (e.g., Eriksson & Strimling, 2015). In the present paper, we will simply assume that all pairs are equally likely to interact.

A Dynamical System

Once the modeler has specified the distribution of different types of individuals, and how types and variants together influence the outcome of social interactions, the evolution of culture in the population can be studied by mathematical means. The trick is to define variables for all entities in one's theory. Our dynamical system is based on the following entities:

- A number of cultural variants. Names for cultural variants will usually be based on the symbol V (for Variant).
- A number of types of individuals. Names for types will usually be based on the symbol T (for Type).
- The probability that a type T individual with cultural variant V will switch to become a V'-bearer when interacting with a type T' individual with cultural variant V'.
- The frequency of different types in the population.
- The frequency of individuals of a specific type of the population who bear a specific cultural variant.
- A series of points in time between which the frequencies of cultural variants change.

When variables are defined for these entities, our assumptions can be translated into a dynamical system. This is a system of equations for how the frequencies of individuals that are of various types, and bear various cultural variants, change from one point in time to the

next. In the supplementary material, this dynamical system is presented as an equation. We shall refer to this dynamical system as our standard model.

It is important to note that our standard model is unlikely to apply to every possible domain of cultural evolution. Researchers may sometimes find it necessary to extend or modify the standard model. Below, we outline a few cases that could require an extension of the model. How these extensions could be implemented is described in the supplementary material.

Change that is not due to social interaction. In the standard model, individuals change their cultural variant only in connection with social interaction with someone who bears a different cultural variant. However, people may also change their cultural variants for other reasons, such as introspection and innovation, memory failure, or because changes in the environment, or in the people themselves, make their current cultural variant unfeasible.

Negative social influence. Our standard model assumes that an interaction between two individuals with the same cultural variant cannot lead to either individual changing. This assumption does not hold in case members of one group (one “type” in our paradigm) desire to be distinct from members of another group. If two members from different groups interact, and both happen to have variant *V*, the result of the interaction could then be that someone changes to another variant. Such an event could be termed *negative social influence*.

A “ladder” of ordered cultural variants. Consider how the habit of cleaning one’s teeth using a toothbrush may vary between individuals. One can clean one’s teeth essentially never, or occasionally, or regularly—perhaps once a day, or twice a day, or even more often. These different habits are cultural variants having a natural ordering (here from less often to more often), which can be thought of as steps on a “cultural ladder” (Eriksson, Enquist, & Ghirlanda, 2007). As in the case of negative social influence above, individuals may then be

assumed sometimes to switch to another cultural variant than the one that was observed. To see why, consider Selda, who seldom cleans her teeth, and Otto, who cleans his teeth twice a day. If Selda happens to discuss teeth hygiene with Otto, she might be influenced to begin cleaning her teeth more often, but not necessarily copying Otto's frequency. Perhaps Selda settles for once a day, perhaps she goes up even higher than Otto to three times a day. Or, if Selda wants to distance herself from Otto, she may stop cleaning her teeth altogether (negative social influence).

Combinations of cultural traits. In human cultural evolution, the spread of cultural traits may depend on the presence of other cultural traits (Enquist, Ghirlanda, & Eriksson, 2011). For example, a belief in man-made climate change may make an individual feel more positively about decreased meat consumption and increased governmental regulation of the energy sector; conversely, an individual who feels strongly about the positives of meat consumption and the negatives of government regulation may be less likely to adopt a belief in man-made climate change. For a discussion of cultural phenomena that can arise when cultural traits are interdependent, see Enquist et al. (2011).

Equilibria and Stability

To discuss the dynamics of a dynamical system properly, it is good to know a few key concepts relating to equilibria and stability. We present these concepts here. Readers who find this section dense can skip over it at first, and refer back to it when necessary. For a thorough introduction to the mathematical study of dynamical systems, see Arrowsmith and Place (1990).

An *equilibrium* is a state that does not change. If the dynamical system reaches an equilibrium, it will remain there. In our standard model, an equilibrium then represents a state

where the distribution of cultural variants in the population has stopped changing. Note that this does not require individuals to stop changing their variants. It only means that the expected number of individuals changing *to* a particular variant is equal to the expected number of individuals changing *from* that same variant.

Our standard model assumes that individuals can only change to cultural variants they observe in an interaction. Under this assumption, no change can occur whenever there is only one cultural variant present in the population. This is known as a *pure equilibrium*, in contrast to a *mixed equilibrium*, in which several cultural variants coexist in the population.

Will a population reach an equilibrium? This question requires the concept of an *asymptotically stable equilibrium*. This is an equilibrium with the property that the process will converge to that equilibrium from any state that is sufficiently close to it. The presence of asymptotically stable equilibria plays a fundamental role for the dynamics that the evolutionary process will exhibit. It is particularly easy to predict the evolutionary process if there is a *unique* asymptotically stable equilibrium that the evolutionary process will converge to from any starting state. This is the case that will arise most often in our examples below. Another case is that there is more than one asymptotically stable equilibrium. We then obtain the important phenomenon of *path dependence*, which means that the outcome of cultural evolution depends on the state in which it starts. A third case is that the system has no asymptotically stable equilibrium. This means that the state will keep changing, which can be called *oscillatory dynamics*.

All of these cases will be illustrated in the catalog of basic cases that we cover below. A summary is given in Table 2. It is possible to simulate a dynamical system with a minimum of programming skills. The graphs we plot below have been obtained from simulations in

Excel. In the supplementary material, we present a mathematical treatment of the standard model, showing how to formally analyze the basic cases presented below.

Table 2. The set of basic cases examined in this paper.

Variants	Types	Asymmetry	Outcome at population level
V and W	Single type	V tends to replace W	V takes over
V1, V2, V3	Single type	V2 tends to replace V1, V3 tends to replace both V2 and V1	V2 may take over temporarily, V3 takes over in the end
V1, V2, V3	Single type	V1 tends to be replaced by V2, which tends to be replaced by V3, which tends to be replaced by V1	Oscillatory dynamics
V and W	Two types of senders	Discriminators send V better, nondiscriminators send V and W equally well	V takes over, equally fast among both types
V and W	Two types of receivers	Discriminators receive V better, nondiscriminators receive V and W equally well	V takes over, faster among discriminators
V and W	Two types of receivers	V-lovers receive V better, V-haters receive W better	Depending on the ratio of V-lovers to V-haters, either a single variant takes over or a mixed equilibrium is approached
V and W	Two types	Asymmetries in both sending and receiving	Path dependence may occur

A Single Type of Individuals

Our main focus will be the important case when there are two cultural variants and at least two types. However, our first set of examples will deal with the special case of *identical* individuals (i.e., a single type), and possibly more than two cultural variants. This case covers

one of the case studies we discussed earlier (Strimling et al., 2018; see also Strimling et al., 2009).

The case of two variants. First, consider the case of only two cultural variants, W and V. In this case, the modeler will only need to specify what tends to happen in an interaction between a W-bearer and a V-bearer: who is most likely to change to the other's variant? Suppose that the modeler specifies that the answer is that it is more likely that V will replace W than vice versa. The resulting dynamical system then specifies that the state in which variant V takes over the entire population is the only asymptotically stable equilibrium. Even if the process starts in a state in which only a minuscule proportion of the population bears variant V, the process will converge toward the state in which everyone bears that variant. If we simulate the process and plot the frequency against time, we obtain the S-shaped curve shown in Figure 2.

Figure 2. An S-shaped curve describing how the frequency of a cultural variant V increases with time (along the horizontal axis) if V tends to replace the alternative variant W in interactions between bearers of each variant. The dashed curve shows the simultaneously decreasing frequency of the alternative variant W.

A succession of cultural variants that are increasingly superior at replacing other variants. In our model of hygiene norms evolution, we considered a case where a succession

of cultural variants is introduced, each new variant eliciting less punishment than the previous ones (Strimling et al., 2018). In the terms of our standard model, this means that we would specify a basic variant V1, a novel variant V2 that tends to replace V1, and an even more novel variant V3 that tends to replace both V1 and V2, etc. The degree of novelty can be represented by the degree of initial rarity in the population. This leads to an evolutionary process in which a new cultural variant temporarily and partially takes over in the population, until overtaken by the next novelty. We illustrate these dynamics with three cultural variants in Figure 3.

Figure 3. A population dominated by variant V1 is invaded by a succession of variants (a small proportion of V2 and an even smaller proportion of V3), each of which tends to replace previous variants in pairwise interactions. V2 will temporarily and partially take over in the population until V3 takes over for good. The horizontal axis represents time.

Many cultural variants, none of which is generally superior at replacing other variants. Another case is obtained if no single cultural variant tends to replace all other cultural variants. For instance, we could have three variants arranged so that V1 tends to be replaced by V2, which tends to be replaced by V3, which tends to be replaced by V1. A mixed equilibrium will then occur for a certain distribution of cultural variants in the population, in which V2's gain from interacting with V1 is exactly counterbalanced by V2's

loss from interacting with V3, etc. However, the mixed equilibrium is not asymptotically stable, nor is any pure equilibrium. The dynamical system will exhibit oscillatory dynamics, in which the cultural variants take turns at dominating in the population. See Figure 4.

Figure 4. Oscillatory population dynamics of three cultural variants arranged so that V1 tends to be replaced by V2, V2 tends to be replaced by V3, and V3 tends to be replaced by V1. The horizontal axis represents time.

Several Types of Senders

So far, we have only dealt with cases of identical individuals. When there are several types of individuals, we need to consider how switching probabilities depend on types. We start by looking at the case when switching probabilities depends on the sender's type, but not on the receiver's type. This case means that some type is better at influencing than another type, at least with respect to some cultural variants. For example, consider our case study of how asymmetry between hygiene-related behaviors in the elicitation of punishment may drive the evolution of hygiene norms (Strimling et al., 2018). In the published model, we assumed no asymmetry between individuals in their use of punishment. However, there are individual differences in how prone people are to punish (Eriksson, Cownden, Ehn, & Strimling, 2014). Should we expect people who are more prone to use punishment to have stricter norms about hygiene and violence? The simplest way to handle individual differences

is to assume that there are two different types: discriminating senders who are better at sending the stricter hygiene norm than the looser hygiene norm (“punishers”); and non-discriminating senders who are equally bad at sending both norms (“non-punishers”). The presence of discriminating senders will ensure that the same norm change as in the original model will occur. The only consequence of including non-discriminating senders is to slow down this norm change; the greater the proportion of non-discriminating senders, the slower norm changes go. See the difference between the left and right panel of Figure 5.

Figure 5 also illustrates another result: Norm change looks exactly the same in the group of punishers as in the group of non-punishers. This is a fundamental theoretical finding of this modeling paradigm: *When the outcome of social interactions does not depend on the type of the receiving individual, but only on the type of the sending individual, different types are not expected to evolve different cultures.*

Figure 5. Cultural evolution of the frequency of an invading cultural variant V when there are two sender types: one type that discriminates between cultural variants by being more likely to send V than other variants, and one type that sends all variants equally well. V then takes over in the population, and at the same rate among discriminators as among non-discriminators (i.e., the curves are identical). Nonetheless, the rate of change increases with a

larger proportion of discriminators in the population: the curve in the left panel, where discriminators are in a 90:10 majority, has a much steeper slope than the curve in the right panel, where the two types are distributed 50:50.

Several Types of Receivers

Next we consider the case when switching probabilities depend on the receiver's type, but not on the sender's type. Given two cultural variants, V and W, we can think of three types of receivers: people who are *more* likely to pick up V than W (*V-lovers*); people who are *less* likely to pick up V than W (*V-haters*); and those that are equally likely to pick up both traits (*non-discriminators*). An example of this case is found in our study of moral opinions (Strimling et al., 2019). Liberals were assumed to be more likely to adopt moral opinions that are supported by arguments based on harm and fairness, whereas conservatives were assumed to not discriminate between different moral arguments. Thus, letting V represent an opinion mainly supported by arguments based on harm and fairness, liberals would be modeled as V-lovers, and conservatives as non-discriminators.

V-lovers and non-discriminators. In a population consisting only of V-lovers and non-discriminators, the only possible outcome of cultural evolution is that V takes over in the entire population. (In other words, the state in which everyone has cultural variant V is a unique asymptotically stable equilibrium.) This can be understood as follows: Because V-lovers are more likely to pick up V than the alternative variant, the frequency of V among V-lovers increases steadily. This increases the probability that a non-discriminator will encounter V, hence also increasing the frequency of V among non-discriminators. As the process is driven by V-lovers discriminating between cultural variants, it also follows that the speed at which variant V takes over will be higher the more discriminating the V-lovers are,

and that the frequency of V is consistently higher among V-lovers than among non-discriminators. These model predictions are illustrated in Figure 6. In our case study of public opinion change across different issues with different connections to moral foundations, these predictions were tested against data (Strimling et al., 2019). The empirical phenomenon in Figure 1 corresponds to the model predictions in Figure 6. Note that in order to match the two figures, we must disregard the two ends of the theoretical curve; we do not find these in empirical data because surveys tend to focus on issues where public opinion is mixed, meaning that support for any opinion is much higher than 0% and well below 100%.

Figure 6. Cultural evolution over time (along the horizontal axis) of the frequency of cultural variant V when there are two receiver types: *non-discriminators* who are equally likely to acquire V and W; and *either V-lovers* (left panel) or *strong V-lovers* (right panel), defined as agents that are either twice or four times likelier to acquire V than to acquire W. In both cases, variant V gradually takes over the entire population, but faster among V-lovers than among non-discriminators. A comparison between the cases shows that the more discriminating the V-lovers are, the faster the cultural change is, even among non-discriminators.

V-lovers and V-haters. Now, consider a population where there are V-lovers and V-haters. The outcome will then depend on how many there are of each type. For certain proportions of the two types, there will be a mixed equilibrium of cultural variants, with a higher frequency of variant V among V-lovers than among V-haters (see Figure 7, left panel). If the proportion of V-lovers is increased, the frequency of V in equilibrium will increase among both types (see Figure 7, middle panel). If the proportion of V-lovers is sufficiently large, variant V will take over the entire population, but faster among V-lovers than among V-haters (see Figure 7, right panel).

Figure 7. Cultural evolution over time (along the horizontal axis) of the frequency of cultural variant V when there are two receiver types: V-lovers (two times likelier to receive V than W) and V-haters (only half as likely to receive V as to receive W). The proportion of V-lovers is set to 50% (left), 60% (middle), or 70% (right). The first two cases yield mixed equilibria, with more V-bearers among V-lovers than among V-haters. As the proportion of V-lovers grows, the equilibrium has more V-bearers among both types. When V-lovers are in a sufficiently large majority, the variant V takes over the entire population.

Differences between senders and between receivers

Finally, consider the general case in which the types of individuals sending and receiving both differ. This case is more complicated to analyze, but some general conclusions

will still hold. Namely, the presence of types that are particularly good at receiving and/or sending a specific cultural variant V will make V more successful, whereas the presence of types that are particularly bad at receiving and/or sending V will make this variant less successful.

Here, we just want to highlight that when types differ among both senders and receivers, the phenomenon of path dependence may arise. Path dependence means that the long-term outcome of cultural evolution will depend on the initial state. Figure 8 illustrates a case where the initial frequencies of cultural variant V within the two types determine which of several asymptotically stable equilibria the population will approach.

Figure 8. Path dependence in cultural evolution over time (along the horizontal axis) of the frequency of cultural variant V (vs. W) in a population with two types. Type 1 is slightly better than type 2 at receiving W , and much worse at receiving V . Moreover, both types receive V better from their own type than from the other type, but receive W better from the other type than from the own type. The only difference between the panels lies in the initial frequencies of variant V . From high initial frequencies, V takes over the entire population (left panel), but from slightly lower initial frequencies, a mixed equilibrium is approached instead (right panel).

Discussion

In this paper, we have argued that emotions play a substantial role in cultural evolution, a role that evolutionary models based on payoffs are poorly equipped to handle. To understand the role of emotions, we think that modelers need to be explicit about the events through which individuals change their cultural variants. In other words, modelers need to consider exactly what the important micro-level mechanisms are. To do this, they need to work with empirical researchers who study such micro-level mechanisms, in particular the emotions triggered by various cultural variants.

By analyzing the cumulative effect of mechanisms operating at the micro-level, it may be possible to understand and predict the major patterns of cultural evolution in a given domain. The trick is to avoid being distracted by the wide range of things that happen at the micro-level, and to focus only on those mechanisms of individual change for which some systematic asymmetry may be found between cultural variants. Emotions are often the reason behind these asymmetries. Two examples were given in the case studies we presented. In our study of hygiene norms, we argued that people's reactions to deviations from their own hygiene norm will differ depending on the direction of the deviation. In our study of moral opinions, we argued that people may feel differently about different kinds of moral arguments. Researchers of emotions can make important contributions by documenting and highlighting such cases of competing cultural variants that are connected to different emotional reactions, either among bearers of the cultural variant, or among observers.

Representation of asymmetries in terms of change probabilities provides the basis of our micro-to-macro theory of cultural change. In particular, this micro-to-macro theory can be expressed in a mathematical model known as a dynamical system. Analysis of the dynamical system yields predictions of how culture will be changing at the population-level.

(Note that it is not possible to conduct the theoretical analysis in the opposite direction, from macro to micro, as many different sets of micro-mechanisms could lead to similar outcomes; see Acerbi, et al. 2016).

A micro-to-macro theory should ideally be validated at both levels; micro and macro. We find the possibility of validating micro-to-macro theories fascinating. Typically, different researchers focus on different levels. They can find more common ground if empirically grounded cultural phenomena on the macro-level can be tied to empirically validated features of micro-interactions through a formal theory that serves to aggregate these micro-interactions to the population-level.

The formal theory will often require only a moderate amount of mathematics. In the present paper, we have shown how to analyze a basic interaction-based dynamical system. We have also discussed how this standard model can be extended to cover additional kinds of change events. A feature of our standard model is the explicit separation of two roles in a social interaction—a “sender” who is potentially influencing a “receiver”—and how these roles may interact with the emotional impact of the cultural variants whose evolution we study. The importance of this separation of roles was illustrated in our analyses. Cultural evolution can be driven both by discriminating senders and discriminating receivers, but these two cases give rise to different dynamics, as culture will typically differ between types that differ as receivers, but not between types that only differ as senders.

The scope and limitations of the modeling paradigm we have presented here should be discussed. For example, can the same modeling paradigm be used for the spread of cultural variants over social media? We believe the answer is yes, because interactions over social media are still interactions with senders and receivers. However, what if people interact not only with each other, but also with some aspect of the environment other than individuals,

such as technologies or institutions? Well, as long as the key source of influence on the environment is the culture of the population at large, it should suffice to extend the paradigm to include state variables of the environment. The researcher must then model how culture influences the environmental state, and how that state in turn influences the culture of the population.

The usefulness of our approach ends when the cultural phenomenon to be studied cannot reasonably be conceptualized as a distribution of cultural variants in a population. Our conceptualization of culture and how it is transmitted is necessarily limited; after all, any model of something as complex as culture will be a huge simplification of reality. The risk of simplification is that something crucial about the phenomenon studied may be missed. On the other hand, the strength of simplification is that it may capture something crucial, and thereby result in a productive way to think about the phenomenon. It is tempting for researchers to try and make models (whether mathematical or verbal) more realistic by incorporating more complexity. We advise caution in this regard, as more complexity seldom yields added insights.

The simplification we make in this paper offers an approach for studying the impact of emotions on cultural change. Our case studies demonstrate that this approach can produce predictions that are non-trivial, testable, and accurate. To make the approach more generally accessible, we have here presented a catalog of basic modeling cases that do not require any further mathematical analysis. We hope this catalog can serve as a toolbox for facilitating and inspiring more research on how emotions drive cultural evolution.

Acknowledgments

This research was supported by the Knut and Alice Wallenberg Foundation [grant number 2015.0005 and grant number 2017.0257].

Declaration of Conflicting Interests

The Authors declare that there is no conflict of interest.

References

- Acerbi, A., Van Leeuwen, E. J., Haun, D. B., & Tennie, C. (2016). Conformity cannot be identified based on population-level signatures. *Scientific Reports*, 6, 36068.
- Arrowsmith, D. K., & Place, C. M. (1990). *An introduction to dynamical systems*. Cambridge University Press.
- Berger, J. (2011). Arousal increases social transmission of information. *Psychological Science*, 22, 891-893.
- Berger, J., & Milkman, K. L. (2012). What makes online content viral?. *Journal of Marketing Research*, 49, 192-205.
- Boyd, R., & Richerson, P. J. (1988). *Culture and the evolutionary process*. University of Chicago press.
- Boyd, R., & Richerson, P. J. (2005). *The origin and evolution of cultures*. Oxford University Press.
- Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution*. Princeton University Press.
- Cownden, D., Eriksson, K., & Strimling, P. (2017). A popular misapplication of evolutionary modeling to the study of human cooperation. *Evolution and Human Behavior*, 38, 421-427.
- De Barra, M., Eriksson, K., & Strimling, P. (2014). How feedback biases give ineffective medical treatments a good reputation. *Journal of Medical Internet Research*, 16, e193.
- Elias, N. (1982). *Power and civility. The civilizing process: Volume II*. Trans. by Edmund Jephcott. New York: Pantheon Books.
- Elias, N., & Jephcott, E. (1978). *The civilizing process* (Vol. 1). Oxford: Blackwell.

- Elster, J. (2015). Explaining social behavior: More nuts and bolts for the social sciences. Cambridge University Press.
- Enquist, M., Ghirlanda, S., & Eriksson, K. (2011). Modelling the evolution and diversity of cumulative culture. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 366, 412-423.
- Eriksson, K. (2012). The nonsense math effect. *Judgment and Decision Making*, 7, 746-749.
- Eriksson, K., & Coultas, J. (2014). Corpses, maggots, poodles and rats: A content bias for disgust in three phases of cultural transmission. *Journal of Cognition and Culture*, 14, 1-26.
- Eriksson, K., Coultas, J., & De Barra, M. (2016). Cross-cultural differences in emotional selection on transmission of information. *Journal of Cognition and Culture*, 16, 122-143.
- Eriksson, K., Cownden, D., Ehn, M., & Strimling, P. (2014). 'Altruistic' and 'antisocial' punishers are one and the same. *Review of Behavioral Economics*, 1, 1-13.
- Eriksson, K., Cownden, D., & Strimling, P. (2017). Social learning may lead to population level conformity without individual level frequency bias. *Scientific Reports*, 7, 17341.
- Eriksson, K., Enquist, M., & Ghirlanda, S. (2007). Critical points in current theory of conformist social learning. *Journal of Evolutionary Psychology*, 5, 67-87.
- Eriksson, K., & Strimling, P. (2015). Group differences in broadness of values may drive dynamics of public opinion on moral issues. *Mathematical Social Sciences*, 77, 1-8.
- Fessler, D. M., Pisor, A. C., & Navarrete, C. D. (2014). Negatively-biased credulity and the cultural evolution of beliefs. *PLOS ONE*, 9, e95167.
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, 96, 1029-1046.

- Haidt, J. (2007). The new synthesis in moral psychology. *science*, 316(5827), 998-1002.
- Heath, C., Bell, C. & Sternberg, E. (2001). Emotional selection in memes: The case of urban legends. *Journal of Personality and Social Psychology*, 81, 1028 – 1041.
- Heuer, F., & Reisberg, D. (1992). Emotion, arousal, and memory for detail. *The handbook of emotion and memory: Research and theory*, 151-180.
- Lang, S. (2012a). *Calculus of several variables*. Springer Science & Business Media.
- Lang, S. (2012b). *Introduction to linear algebra*. Springer Science & Business Media.
- Lumsden, C. J., & Wilson, E. O. (1981). Genes, mind, and culture. Harvard University Press.
- Miton, H., Claidière, N., & Mercier, H. (2015). Universal cognitive mechanisms explain the cultural success of bloodletting. *Evolution and Human Behavior*, 36(4), 303-312.
- Nichols, S. (2002). On the genealogy of norms: A case for the role of emotion in cultural evolution. *Philosophy of Science*, 69, 234-255.
- Peters, K., Kashima, Y., & Clark, A. (2009). Talking about others: Emotionality and the dissemination of social information. *European Journal of Social Psychology*, 39, 207-222.
- Pinker, S. (2011). *The better angels of our nature: The decline of violence in history and its causes*. Penguin.
- Rabinovich, Y., Sinclair, A., & Wigderson, A. (1992). Quadratic dynamical systems. In *Foundations of Computer Science, 1992. Proceedings., 33rd Annual Symposium on* (pp. 304-313). IEEE.
- Richerson, P., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.
- Sandholm, W. H. (2009). Evolutionary game theory. *Encyclopedia of Complexity and Systems Science*, 3176-3205.

- Schelling, T. C. (2006). *Micromotives and macrobehavior*. WW Norton & Company.
- Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge University Press.
- Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Cambridge, MA: Cambridge.
- Strimling, P., De Barra, M. & Eriksson, K. (2018). Asymmetries in punishment propensity may drive the civilizing process. *Nature Human Behaviour*, 2, 148-155.
- Strimling, P., Enquist, M., & Eriksson, K. (2009). Repeated learning makes cultural evolution unique, *Proceedings of the National Academy of Sciences*, 106, 13870-13874.
- Strimling, P., Vartanova, I., Jansson, F., & Eriksson, K. (2019). The connection between moral positions and moral arguments drives opinion change. *Nature Human Behaviour*, in press.