

Research



Cite this article: Hébert-Dufresne L, Waring TM, St-Onge G, Niles MT, Kati Corlew L, Dube MP, Miller SJ, Gotelli NJ, McGill BJ. 2022 Source-sink behavioural dynamics limit institutional evolution in a group-structured society. *R. Soc. Open Sci.* **9**: 211743.

<https://doi.org/10.1098/rsos.211743>

Received: 9 November 2021

Accepted: 4 February 2022

Subject Category:

Science, society and policy

Subject Areas:

behaviour/evolution/theoretical biology

Keywords:

source-sink dynamics, institutions, behavioural diffusion, cultural evolution, cooperation

Author for correspondence:

Timothy M. Waring

e-mail: timothy.waring@maine.edu

[†]These two authors contributed equally.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5882110>.

Source-sink behavioural dynamics limit institutional evolution in a group-structured society

Laurent Hébert-Dufresne^{1,2,3}, Timothy M. Waring^{6,7,†}, Guillaume St-Onge^{5,9}, Meredith T. Niles³, Laura Kati Corlew¹⁰, Matthew P. Dube¹¹, Stephanie J. Miller^{7,8}, Nicholas J. Gotelli⁴ and Brian J. McGill^{7,8}

¹Department of Computer Science, ²Vermont Complex Systems Center, ³Department of Nutrition and Food Sciences, and ⁴Department of Biology, University of Vermont, Burlington VT, USA

⁵Département de physique, de génie physique et d'optique, Université Laval, Québec (Québec), Canada G1V 0A6

⁶School of Economics, ⁷Mitchell Center for Sustainability Solutions, and ⁸Mitchell Center for Sustainability Solutions, University of Maine, Orono ME, USA

⁹Centre interdisciplinaire en modélisation mathématique, Université Laval, Québec (Québec), Canada G1V 0A6

¹⁰Department of Social Science, and ¹¹Department of Computer Information Systems, University of Maine at Augusta, Bangor ME, USA

LH-D, 0000-0002-0008-3673; TMW, 0000-0001-7364-1130; GS-O, 0000-0001-5415-3748

Social change in any society entails changes in both behaviours and institutions. We model a group-structured society in which the transmission of individual behaviour occurs in parallel with the selection of group-level institutions. We consider a cooperative behaviour that generates collective benefits for groups but does not spread between individuals on its own. Groups exhibit institutions that increase the diffusion of the behaviour within the group, but also incur a group cost. Groups adopt institutions in proportion to their fitness. Finally, the behaviour may also spread globally. We find that behaviour and institutions can be mutually reinforcing. But the model also generates behavioural source-sink dynamics when behaviour generated in institutionalized groups spreads to non-institutionalized groups and boosts their fitness. Consequently, the global diffusion of group-beneficial behaviour creates a pattern of institutional free-riding that limits the evolution of group-beneficial institutions. Our model suggests that, in a group-structured society, large-scale

1. Introduction

There is broad agreement that global crises such as climate change and pandemic response will require society to change. Therefore, a robust science of social change is needed. It has recently been suggested that the study of collective behaviour should be considered a crisis discipline [1]. Recent research, including the study of social tipping points [2], their impact on policy efficacy [3,4], and the feedbacks between policy and preferences [5], suggests a science of social change may be possible. But what constitutes social change? At a minimum, we suggest that social change must include both change in individual behaviours and change in institutions. Here, we study a model of social change that couples behavioural change with policy evolution.

The most difficult societal problems require both widespread change in individual behaviour and institutional change. But individual behaviour and group-level institutions are mutually interdependent, making solutions difficult to describe and achieve. On the one hand, large-scale behavioural change is slow or unpredictable unless supported by group-level institutions and policy. On the other hand, policy changes often follow from broad changes in individual preferences, beliefs or behaviour (e.g. gay marriage, civil rights). For example, no serious proposals for addressing climate change rely on the spread of voluntary individual behaviour alone; institutional support such as requirements and enforcement are clearly required. Yet without widespread individual support for climate policy, nations are unlikely to make institutional and structural change. This reciprocal causation makes addressing major societal goals particularly challenging because there the two most common approaches focus on only one aspect of the system. A policy-first approach uses a top-down change in institutions to affect large-scale behavioural change, for example by requiring and enforcing a behaviour (e.g. speeding fines, tax incentives, smoking bans). Alternatively a behaviour-first approach seeks to encourage voluntary behavioural change through bottom-up mechanisms such as unenforced behavioural standards (e.g. social distancing guidelines, voluntary participation in environmental programmes, and public service announcements generally). Both approaches have benefits and drawbacks.

Policy analysis is complex and policy outcomes are heavily determined by contextual factors. Policy design must therefore focus on local impacts, and thus often ignores where institutions come from, both in terms of how they originate via the policy process [6], and in terms of how policies spread via policy diffusion [7,8]. Institutions, governance, laws and rules emerge from the complex inter-individual process of cultural evolution and self-organization within human groups [9,10]. As a result, institutions are themselves distributions of individual cultural traits and behaviours [11]. However, research on the efficacy of policy options, e.g. [12–14], tends to be isolated from research on policy diffusion [7,15] and from research on the evolution of institutions [9,16–18]. Research that does link behavioural change to policy change typically focuses on only one half of the causal system. For example, a policy analyst might study how new COVID-19 policies determined individual mask-wearing behaviour. Conversely, a social scientist might study the spread of vaping behaviour through a social network and its impact on institutional responses. These are useful and necessary approaches; but behaviour and institutions are strongly reciprocally linked, as summarized in figure 1.

By contrast, a behaviour-first approach relies on the endogenous processes of behavioural adoption and transmission. Behaviour-first strategies, such as nudges [19] and viral advertising efforts, seek to leverage endogenous behavioural change for certain outcomes. Research on social tipping points shows that social momentum can be harnessed to improve well-being by boosting the spread of beneficial social norms and behaviours [2]. Similar research shows that these social tipping points can be caused by conformity and behavioural transmission and can create behavioural spillovers that determine the efficacy of policy [3,4,10]. When successful, behaviour-first approaches such as voluntary guidelines can lead to the diffusion of behaviour within communities. Research of this sort typically seeks to find a social tipping point past which the beneficial behaviour becomes ubiquitous [20]. If the behaviour becomes associated with values or social identities [21], the result can be the evolution of a self-reinforcing system of social norms: a new standard of behaviour [5,22]. However, behavioural diffusion and viral campaigns do not reliably generate social-tipping points, especially when the behaviour is costly.

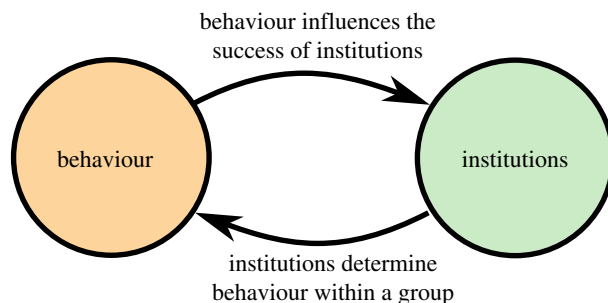


Figure 1. Individual behaviour and group institutions form a dynamical system with reciprocal causation. To understand the endogenous nature of social change, we must model the causal feedbacks between institutions and behaviour, and to influence social change we must be able to find and move towards beneficial states that are self-reinforcing.

A policy-first approach promotes a given behaviour via a change in institutions such as laws, rules or other formal group-level procedures. These institutional changes may simply be rules promoting or requiring behaviour, or they may entail behavioural support or enforcement. A policy approach can achieve rapid and large-scale behavioural change by altering the choices faced by an entire group. The scale and strength of the policy can determine the scale of adoption of the target behaviour to a great extent. However, incentives can undermine intrinsic motivation for the behaviour [23–25]. Moreover, behavioural requirements can be perceived as coercive even when they result from collective democratic policymaking, as has been evidenced with the COVID-19 pandemic [5]. Consequently, policy based on behavioural requirements can fail if people stop participating when the enforcement or incentive is withdrawn. Thus, policy-first solutions are often powerful but brittle.

One critical aspect of social change that is often overlooked are the feedbacks between individual behaviour and group-level policies [22,26]. There are multiple ways in which policies, whether voluntary or coercive, can cause self-defeating social dynamics among citizens and ‘crowd out’ beneficial behaviour. Those reactions depend upon the culture of the group, which in turn has been conditioned by previous policies and institutions [5]. Therefore, it is not possible to evaluate policy efficacy without understanding its reciprocal interactions with individual traits and behaviour.

We think it is useful to consider behavioural transmission and institutional change as entangled parts of a larger process of social change. Social change is attractive as a policy framework because it combines the benefits of endogenous (bottom-up) behavioural change and (top-down) interventions and policy, and allows us to ask when and how policy and behaviour come to be mutually reinforcing. However, unlike traditional policy frameworks, social change is at least partly an endogenous process, and cannot simply be implemented. In this paper, we ask how social change occurs when behaviours and institutions are mutually determined. We present a complex mathematical model for studying endogenous social change that couples simple mechanisms for the evolution of institutional policies and for the transmission of individual behaviours.

2. Linking behaviours and institutions

Behaviours and institutions are similar phenomena on different social scales; both are cultural traits. Behaviour, in most cases, is an individual cultural trait, or the expression of one, and institutions are group-level cultural traits. Both are invented, learned, modified, copied and re-transmitted, diffusing and evolving within a larger population. Thus, the simplest method to unify behavioural change and institutional change is to consider their transmission in parallel [18].

2.1. The transmission of behaviour

Behavioural diffusion has been studied from multiple perspectives, including the diffusion of innovation [27], behavioural contagion [28,29] and cultural transmission [30]. The transmission of cultural traits via social learning, imitation and teaching is central to the process of cultural evolution, and helps drive linguistic change, economic development and technological accumulation [31–33]. Other cultural traits including the language, preferences, identities, beliefs and values that determine behaviour are also transmitted via social learning between individuals [34]. This social science is complemented

by research in disease modelling on behaviours that respond to, and spread with, an infectious disease [35,36].

Cooperative and pro-social behaviours are especially important, underlying successful group-level collective action by producing benefits beyond the actor. Here, we consider cooperative behaviour to be behaviour that carries group-level benefits. We are particularly interested in how institutions can promote the simple diffusion of such behaviours, which is motivated by previous results using a wealth of different approaches. Behavioural science shows that humans are conditional cooperators [37], cooperating when either the situation [38] or the social institutions support cooperative behaviour [39]. Moreover, cooperative and pro-social behaviour may sometimes be culturally transmitted between individuals within and between groups [40–42]. However, because cooperative behaviours often come at a cost to the actor, they rarely persist or spread on their own. The factors that promote the evolution of cooperation have also been heavily studied. These mechanisms have two common features: they are payoff-modifying in that they make cooperation less costly or non-cooperation more costly, and group-focused in that they operate within some group of individuals somehow defined. Cooperation stabilizing factors include decentralized individual-level mechanisms to exclude non-cooperators (e.g. ostracism) or reduce their payoffs (e.g. peer punishment), and centralized group-level institutions (e.g. pooled-punishment, or policing) which also change payoffs.

Since cooperation grows best in groups, inter-group processes such as competition and selection typically spur the evolution of cooperation in certain regards. Cooperation and institutional traits that support it can spread when human groups compete [18,43–45]. For example, trust tends to grow to higher levels in more competitive industries [46]. Game theoretic and evolutionary models show how cooperation can spread through group-level selection even when individuals face a social dilemma with other group members [47–49]. Analysis of these models reveals that a certain critical threshold of group-level pressure is required to make cooperation evolve generally. This cultural group selection mechanism helps explain both the spread of cooperation and the emergence of institutions that support it [18,26,48,50].

2.2. The evolution of institutions

Institutions, including centralized mechanisms for punishment, redistribution, division of labour and collective decision-making are a type of group-level cultural trait [11]. In comparison to decentralized behavioural regimes such as peer punishment or ostracism, institutions are widely considered the most effective and important mechanisms by which human groups stabilize cooperative behaviour. Institutions can stabilize cooperation and drive widespread behavioural adoption because they can render a cooperative behaviour individually beneficial by changing the outcomes and payoffs from each behavioural option. Institutional evolution is therefore believed to gradually increase the functionality of institutional arrangements for the groups that adopt them. For example, the most studied institutional features, Ostrom's (1990) institutional rules [51], are considered to be group-beneficial cultural adaptations which evolved in large part to help groups maintain cooperation in important domains of shared interest [52]. Thus it is believed that institutions emerge and spread largely because they enable groups to resolve collective action problem and social dilemmas [33,44]. High levels of institutionalization, such as are found in rich industrial societies, can support a cooperative and productive economic system.

Institutions also spread between groups, via processes largely parallel to the mechanisms of social learning at the individual level. Research on policy diffusion shows that the spread of policy, rules and institutions between groups is also an important determinant of change [15,53]. Institutions may spread between companies, nations, municipalities, sports teams, or in any population of human groups. So, naturally, institutional evolution depends on the transmission of institutions between groups [9]. Thus, when organizations, nations and companies may imitate, modify and re-transmit helpful institutions, the result is the evolution of organizational structure, business models, municipal rules and national law. In fact, the concurrent spread of behaviour change and institutional support was recently identified as a needed avenue of research in public health modelling [54]. But currently, the transmission of individual-level behaviour and group-level institutions are typically considered in isolation.

Therefore, the relationship between the diffusion of behaviour and the spread of group policies and institutions is important for societal outcomes, yet remains understudied. On one hand, cooperative behaviour only tends to spread when institutions support it and render it less costly or altruistic for participants. On the other, institutions themselves only emerge when a group has a

shared identity and intention to resolve a common goal: when cooperation is not a problem. So, the cooperation-institutions issue is a chicken-and-egg question; to understand one, we must understand both.

To help consider how policy and behaviour can produce beneficial societal outcomes, we ask how behaviour and institutions interact in a group-structured population. We develop a model of the coupled diffusion of individual behaviour and group-level institutions using a novel approach that combines epidemiological (e.g. [55]) and evolutionary approaches (e.g. [50]). We are interested in a group-beneficial behaviour, such as a cooperative or pro-social behaviour which will tend to dissipate without group-level support, but that can spread when a group adopts a policy to promote it. As our goal is to explain social change, not cooperation *per se*, we do not model individual costs and benefits in either evolutionary or economic terms, nor do we include utility maximization or payoff-biased imitation. We only model a group-beneficial behaviour that diffuses poorly on its own. The greater the group's efforts to promote the behaviour, the more rapidly it spreads within that group. We couple the diffusion of this behaviour with a process of group-level institutional evolution, in which groups with more of the behaviour achieve better group-level outcomes. Groups with better outcomes are preferentially imitated, leading to the spread of institutions. We explore how this coupled system of behavioural and institutional evolution operates as a whole using a model that combines epidemiological spread of individual behaviour with group-level selection of institutions. Model code in C++ is available <https://github.com/LaurentHebert/group-cultural-adaptation>.

3. Model

We assume a population of size $N \gg 1$, divided into M groups of equal size $n = N/M$. Groups contain i adopters of the behaviour and $(n - i)$ non-adopters. The natural diffusion rate, β , is the rate at which the behaviour spreads from adopters to non-adopters within groups at a basic institutional level that allows but neither promotes nor discourages the behaviour. Adopters also lose that behavioural trait at a rate γ . Groups adopt an institution from a series of discrete levels of institutional strength, ℓ , which modify the rate of behavioural diffusion to be $\ell\beta$. Note that without group-level institutions—i.e. at $\ell = 0$ —the behaviour does not spread at all. Therefore, like behaviours that benefit the group but incur individual costs such as cooperation and altruism, the behaviour we model will dissipate if not also adopted by one's social contacts. The institutional strength of a group can therefore be conceptualized as setting the susceptibility of individuals in that group to adopting the behaviour. Finally, global behavioural diffusion, ρ , allows the behaviour to spread between groups, based on the institutional strength of the receiving group.

We then implement this system of behavioural diffusion within and among groups through a set of master equations [56] that tracks the fraction of groups $G_{i,\ell}(t)$ with i adopters and institutional level ℓ . Omitting the time dependency for simplicity, the behaviour diffuses following:

$$\begin{aligned} \frac{d}{dt} G_{i,\ell}^{\text{diff}} = & \ell\beta[(i - 1) + R](n - i + 1)G_{i-1,\ell} \\ & - \ell\beta(i + R)(n - i)G_{i,\ell} \\ & + \gamma(i + 1)G_{i+1,\ell} - \gamma i G_{i,\ell}. \end{aligned} \quad (3.1)$$

In this equation, the first term corresponds to diffusion events bringing groups from state $(i - 1, \ell)$ to state (i, ℓ) , which occurs proportionally to the internal diffusion rate $\ell\beta$ times the number of non-adopters $n - (i - 1)$ who could adopt the behaviour. The factor in square brackets corresponds to the number of all adopters with influence over the non-adopters within that group, which includes $i - 1$ adopters within the group and the non-adopters in other groups who are exposed via global diffusion, i.e.

$$R = \rho \sum_{i',\ell'} i' G_{i',\ell'}, \quad (3.2)$$

where primes simply denote variables over which we sum to calculate a global quantity.

The second term in the master equation also corresponds to diffusion events but now taking groups out of state (i, ℓ) and into state $(i + 1, \ell)$. Finally, the last two terms correspond to relaxation events where adopters revert to their old behaviour; bringing groups from state $(i + 1, \ell)$ to (i, ℓ) or from (i, ℓ) to $(i - 1, \ell)$.

Additionally, we model institutional evolution among groups. Institutions are defined as a series of discrete levels of institutional strength $\ell \in [\ell_{\min}, \ell_{\max}]$, here set as $\ell \in [0, 5]$ but easily generalizable.

Greater levels of institutional strength come with greater efficacy in promoting the group-beneficial behaviour (i.e. $\ell\beta$) and with greater effort. Groups pay a cost c for each additional level of institutional strength, and gain a collective benefit b for each adopter in the group. Institutions evolve via a group selection process in which groups select their institutional strategy based upon the payoffs to groups with other strategies. We assume that a given level of institutional strength ℓ has a perceived fitness Z_ℓ given by

$$Z_\ell = \frac{\sum_i \exp(bi' - c\ell)G_{i,\ell}}{\sum_i G_{i,\ell}}, \quad (3.3)$$

where the denominator acts as a normalization factor. Note that only the fitness of a group's strategy is directly observed and not the individual behaviour of its members. Based on this perceived fitness, groups can change their institutional strength following a second master equation,

$$\begin{aligned} \frac{d}{dt} G_{i,\ell}^{\text{select}} = & \rho[G_{i,\ell-1}(Z_\ell Z_{\ell-1}^{-1} + \mu) + G_{i,\ell+1}(Z_\ell Z_{\ell+1}^{-1} + \mu)] \\ & - \rho(Z_{\ell-1} Z_\ell^{-1} + Z_{\ell+1} Z_\ell^{-1} + 2\mu)G_{i,\ell}. \end{aligned} \quad (3.4)$$

These terms track flows of groups from institutional strength ℓ to $\ell+1$ and $\ell-1$ (and vice versa). These flows are assumed to occur proportionally to the relative fitness of different levels (e.g. a group moves from ℓ to $\ell+1$ proportionally to $Z_{\ell+1}/Z_\ell$) and the rate of global behavioural diffusion ρ , which is equivalent for individual behaviour and group institutions. Importantly, we also add a constant rate of transition μ regardless of fitness which allows the invention of new institutional levels (e.g. discovery of unoccupied institutional strategies).

The total dynamics of our model is then given by a set of master equations,

$$\frac{d}{dt} G_{i,\ell} = \frac{d}{dt} G_{i,\ell}^{\text{diff}} + \frac{d}{dt} G_{i,\ell}^{\text{select}}, \quad (3.5)$$

defined over $i \in [0, n]$ and a set of integers for levels of institutional strength $\ell \in [\ell_{\min}, \ell_{\max}]$. We can thus track the dynamics of our system by numerically integrating all equations starting from arbitrary initial conditions (constrained only by $\sum_{i,\ell} G_{i,\ell} = 1$).

The system is sufficiently complex to rule out analytical descriptions of equilibria, but we explore stability conditions analytically in electronic supplementary material, appendix. We simulate systems with identical initial conditions, specifically a uniform distribution of groups across institutional levels and a binomial distribution of i with average of 1% adopters. Stable long-run values were recorded when $t \geq 10\,000$ and if the expected average difference Δi in groups was less than 10^{-10} over a $\Delta t = 1$.

4. Results

The model produces a set of coevolutionary dynamics between behaviours and institutions, some of which were unexpected, and hold value as tools for understanding social change. We summarize model behaviour in figures 2–5.

As expected, the model confirms that cultural group selection on institutions can drive evolution of cooperation or pro-social behaviour, as others have demonstrated [26,47,48]. We find that for behavioural-institutional coevolution to take off, two fundamental criteria must be met: institutional implementation costs, c , must be below a threshold value (see figure 2, and the rate of behavioural transmission within groups, β , must be above a critical threshold (figure 3). These thresholds are sharp. Even small changes in c and β can have catastrophic effects (figure 3). In fact, we find a rich mix of both continuous and discontinuous emergence of the behaviour depending on institutional levels and parameter values. We explore the mathematical conditions around these transitions in electronic supplementary material, appendix.

Qualitatively, no institutions are possible if institutional costs are too high, and the behaviour never spreads. Lowering institutional costs slightly below the threshold results in a sharp discontinuity such that institutions and behaviour both flourish. Further decreases do not change qualitative outcomes. A similar threshold also exists for within-group behavioural transmission, β . Unlike the cost threshold, however, increases in β above the threshold result in further spread of both behaviour and institutions. However, if costs are low, institutions may still evolve even with a very low β (figure 4a).

Naturally, simple institutions (i.e. $\ell = 1$) emerge and become most common across a large range of parameters (c, β, ρ) (figure 4c). Stronger institutions (i.e. $\ell \geq 2$) also evolve if costs are sufficiently low

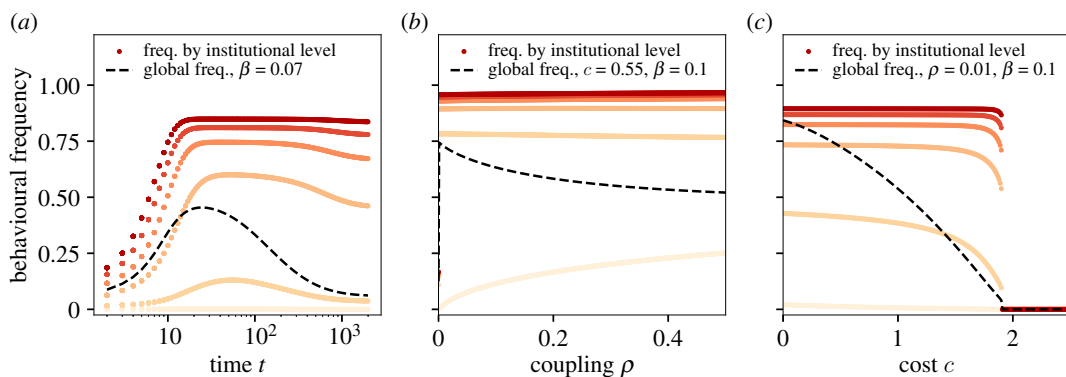


Figure 2. Frequency of behaviour in groups with different institutional strength. Within groups, the frequency of cooperative behaviour follows the strength of institutions, shown in shades of red ($\ell = 0$ in light beige and $\ell = 5$ in dark red). (a) The time dynamics of global behavioural frequency (black dashed line) and behaviour in groups can include patterns of surge and collapse. (b) Increasing global diffusion of behaviour, ρ , does not affect the frequency of behaviour in institutionalized groups but increases adoption of the behaviour in groups that have no institution (light beige). This shift increases the relative fitness of weaker institutions, leading to a decrease in global adoption of cooperation (figure 4). (c) Increasing the cost of institutions steadily decreases the global frequency of behaviour, eventually reaching an inflection point (here 1.9) above which increasing costs causes synchronized discontinuous collapse of cooperation across all groups. Other parameters are fixed at $n = 20$, $\gamma = 1.0$, $b = 0.18$ and $\mu = 10^{-4}$.

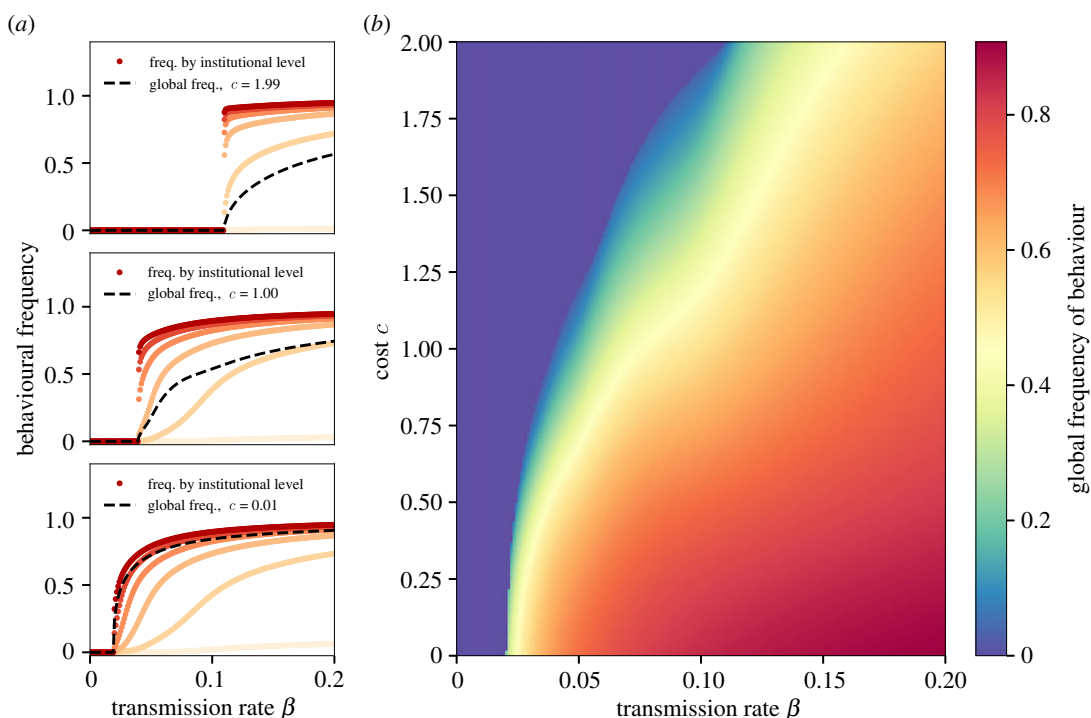


Figure 3. Behavioural transmission and institutional cost drive behaviour frequency. Institutional costs ($c = 0.01, 1.0, 1.99$) interact with behavioural transmission β to determine frequency of cooperation behaviour. (a) At lower costs behaviour emerges gradually with increasing transmission, but at higher costs this occurs more rapidly and uniformly across groups at different institutional levels. In groups with institutions $\ell \geq 2$, behaviour spreads extremely rapidly, while groups with $\ell = 1$ see a slower and continuous activation with increasing β . Groups without institutions, $\ell = 0$, displayed as the lightest shade, cannot generate cooperation but still see an increase in cooperation due to the global transmission from other groups. (b) This same interaction can be seen in the global frequency of behaviour across all combinations of β and c . Parameters fixed at $n = 20$, $\gamma = 1.0$, $b = 0.18$ and $\mu = 10^{-4}$.

(figure 4). However, as costs rise towards the threshold, the greatest level of institution that can be sustained declines steadily. Note that our model is harsh on the evolution of stronger institutions, because we do not allow institutions to modify costs, which they are typically assumed to do. Therefore, we suspect that the model underestimates institutional evolution.

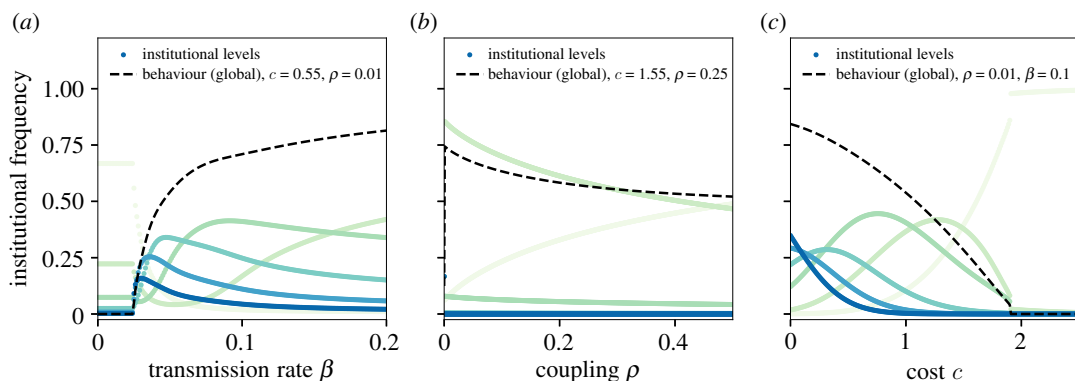


Figure 4. Institutional frequency among groups. The adoption of different levels of institution (i.e. $\sum_i G_{i,\ell}$) is shown with $\ell = 0$ in light green and $\ell = 5$ in dark blue. (a) At low values of β , groups with $\ell = 0$ are most common. As the transmission rate crosses the threshold, stronger institutions ($\ell > 2$) proliferate rapidly, while weaker institutions first decline in frequency, then grow to become dominant. Institutions of $\ell = 2$ or $\ell = 3$ become most popular at intermediate values of β . Stronger institutions are progressively favoured at lower value of cost c and global diffusion ρ . (b) Increasing the global diffusion rate ρ confirms the emergent institutional free-riding problem: the no-institutions strategy progressively gains in popularity as global diffusion of cooperation increases, eventually becoming dominant. (c) Stronger institutions are favoured at lower cost. As institutional costs increase, lower levels of institution are progressively preferred. This dynamic continues until the lowest non-zero institutional level $\ell = 1$ also loses its fitness at the cost threshold, where these groups experience a discontinuous transition. Parameters fixed at $n = 20$, $\gamma = 1.0$, $b = 0.18$ and $\mu = 10^{-4}$.

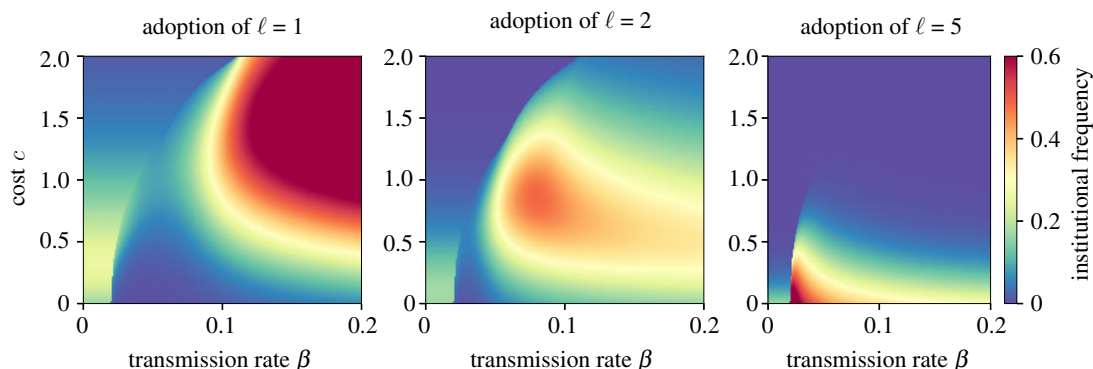


Figure 5. The prevalence of institutions of different strengths varies by costs and transmission rate. Panel replicates the numerical experiment of figure 3, but shows the adoption of institutions of level 1, 2 and 5, respectively. This comparison highlights the phenomenon of institutional localization in which a given institutional level dominates the fitness landscape in some subset of parameter space. This phenomenon is analytically explored in electronic supplementary material, appendix.

Under certain conditions, the model exhibits an unexpected pattern of bi-modal institutional success. This occurs when a certain level of institutionalization is required to maintain the behaviour locally, but only some groups become self-sustaining. For example, close to the sustainability thresholds, it is possible to arrive at a state in which approximately 20% of groups have no institution, approximately 25% with level 3 institutions, but very few have level 1 or 2 institutions (figure 4a). This state emerges when the intervening institutional level ($\ell = 1$ or 2) has lower fitness than either $\ell = 0$ or $\ell = 3$, hindering non-institutional groups from moving to stronger institutions. We call this phenomenon institutional localization. We explore it visually in figure 5 and analytically in details in electronic supplementary material, appendix.

The rate of global behavioural diffusion, ρ , also creates a counterintuitive effect. As ρ increases, total behavioural adoption and institutions both *decrease* (figure 4b). Increased global diffusion allows the cooperation to spread into groups that do not have institutions to generate or support it. Cooperation in receiving groups does not persist, but still increases their fitness. Higher fitness outcomes for poorly institutionalized groups then creates selection against stronger institutions. In essence, high between-group transmission, ρ , creates *behavioural source-sink dynamics*: behaviour generated in a group with institutional support diffuses to other groups, improving the outcomes of those groups

without effective local institutions. Put in evolutionary terms, this ‘institutional free-riding’ caused by between-group behavioural transmission can halt cultural group selection and the evolution of group-beneficial institutions.

The source-sink dynamics of cooperation have an important influence on model outcomes. In the region around, the emergence threshold model behaviour is complex and outcomes are sensitive to initial conditions. In this region, we find non-monotonic time series in which behaviour and institutions can spread together for a time, only to collapse later (figure 2a). In certain parameter combinations, the model exhibits multiple stable equilibria, ending either in the collapse of behaviour and institutions (e.g. all groups in $i = \ell = 0$) or with a steady-state non-zero level of both behaviour and institutions, depending finely on initial conditions. This model thus provide an interesting mathematical sandbox for future work considering how to design promotion of beneficial collective behaviour and institutional support.

5. Conclusion

Our model helps us understand social change when behaviour and institutions are mutually linked. Connecting the diffusion of group-beneficial behaviour to the spread of policy and supporting institutions reveals a set of social dynamics that are not often studied. Our analysis shows that group-beneficial behaviour can be subject to source-sink dynamics in which cooperation generated by groups with strong institutions spreads to less-institutionalized groups. These behavioural source-sink dynamics create an emergent pattern of institutional free-riding in which receiving groups benefit from the imported cooperative behaviour, but cooperation dies away in those groups without institutional support. Thus, receiving groups act as a cooperation sink. We also find that when the global spread of cooperation, ρ , is sufficiently strong, these dynamics can halt the evolution of institutions. These dynamics can dampen or halt cultural group selection for institutions when the behavioural diffusion increases the fitness of receiving groups with weak institutions.

Source-sink dynamics of cooperation may be influencing social evolution all the time, and examples abound. One area of interest is when the provision of public goods is difficult to understand or perceive. When the cooperation of one group creates benefits for other groups, institutional free-riding may result. One example of this may occur in national tax systems when tax revenue is shared between regions. If such revenue sharing enhances the perceived success of recipient states, it may also undermine institutional evolution among states. Another immediate example concerns the heterogeneous and slow spread of robust public health policy for fighting COVID-19 in the USA [57–59]. States with robust policy (i.e. masking, social distancing and vaccination requirements) probably spread those beneficial behaviours to states with weak public health policy. This behavioural spillover could artificially increase the actual or perceived success of states with weak policies, slowing the uptake of effective policies around the country.

Source-sink dynamics of cooperation might also help explain long-term patterns of social change, such as the decline of unions in the USA [60]. Collective bargaining often improves work conditions and pay for union members, but can also benefit non-union members. For example, in right-to-work states, people can reap the rewards of a union without joining [61]. This undercuts the strength of the local union, but it may also reduce union fitness generally, slowing their spread elsewhere. Thus, limiting negotiated benefits to members might improve the correlation between institution and behaviour, helping to maintain both. Another example comes from food safety regulations. A lack of awareness of the role of the safety regulations can lead municipalities to repeal them. But, when people mostly consume food from safely regulated production systems, those non-regulated municipalities are protected from negative consequences, and the reduced safety standards can spread. Something similar has been observed in Maine, where reduced food safety standards have spread between many small towns [62]. These examples suggest that policies that create public goods can be made more durable if they also make their supply of those public goods highly observable. Our impetus for this project is the category of societal or global-scale problems such as climate change which may require large-scale social change. We believe the source-sink dynamics of group-beneficial behaviour provide a new and useful framework and are worthy of additional study.

In combining behavioural transmission with institutional evolution, the model shows that it is possible to overcome limitations of the policy-first and behaviour-first approaches. Specifically, our model has implications for intervention efforts to achieve large societal goals that require social change which do not emerge from the standard approaches. First, we find that, despite reciprocal

dynamics between behaviour and policy, the spread of institutions has a greater influence on behavioural diffusion than the reverse. This is an important insight for both social scientists and policymakers in the age of nudges and viral behavioural campaigns: as a group-level cultural trait, policy has an outsized influence on the evolution of a group-structured society. But social change is not owned and cannot be controlled by any one actor, even including legitimate governments. Any actor within society might look to influence social change. Indeed, beyond political domains many actors already intentionally and strategically influence social change. For example, for-profit corporations spend large sums of money in advertisements to modify consumer preferences and purchasing behaviour. Therefore, it is useful to provide intervention points for actors in different social positions. We focus on two levels, first the level of the group (such as a state or organization) and second, the level of the population of groups (such as a nation).

Group-level actors, such as state governments or organizations, have two major avenues for action that emerge from our framework. First, groups can work to improve the spread of the cooperative or beneficial behaviour within the group. For example, create meaningful rewards for the group-beneficial behaviour, or ways of celebrating the contributions of those who contribute. Second, groups can adopt new institutions (which enhance behavioural transmission) by learning from other innovative groups. Innovative institutions may be social in nature, such as finding and enhancing ways for individuals to reward, recognize and congratulate each other on their contributions. But groups should also be careful to avoid relying on externally generated cooperation, lest it cease to flow. This framework could be useful for economic development in rural states.

Population-level actors, such as national governments, can work to improve the spread of policies that support beneficial behaviour between groups. There are two main fronts in this effort. First, reducing the costs of implementing policies that support group-beneficial behaviour can be very effective at spreading both behaviour and policy. For example, a central government could provide financial support to cities that adopt a policy to support climate-friendly behaviour. Second, increasing between-group learning of policy options can accelerate behaviour-policy coevolution. For example, the US Federal Government could enhance the transmission of beneficial state policies by supporting conferences between state governments to facilitate the exchange of ideas and solutions in a domain of interest (e.g. combating addiction).

Actors at the population level must also contend with source-sink dynamics of behaviour which can undermine cooperation and institutional evolution. This process is determined in part by the rate of transmission of group-beneficial behaviour between groups. Therefore, our model suggests that efforts to spread cooperation globally are likely to be somewhat self-defeating. Furthermore, we would caution against strategies for reducing the transmission of cooperation between groups, as such policies could counteract the transmission of cooperation *within* groups, which is of even greater importance. Instead, population-level actors should work to create a pattern of reinforcement between group-beneficial behaviour and local institutions. One way to accomplish this is through correlated pilot projects: efforts to boost the *correlated* spread of beneficial behaviour and supporting institutions. In our model, behaviour and institutions spread best when they are correlated across groups and with beneficial outcomes. Therefore, policy approaches might focus on promoting strong pilot studies and local public trials with high chances of success to be copied elsewhere. Future research could use this model to study examples of pilot project diffusion. For example, a study of coastal management pilot projects in South Africa suggests that the coupled diffusion of cooperation and institutions may be useful for achieving societal transitions to more sustainable regimes [63]. It seems likely that similar dynamics might underlie the diffusion of environmental regulation [15,64,65].

Our approach can be improved and expanded in a number of useful directions. First, the current model ignores individual payoffs and decision-making. Adding these would bring it in line with the game theoretic and evolutionary literature on cooperation, and might have implications for the extent of source-sink dynamics of group-beneficial behaviour. For example, one could investigate new mechanisms to accelerate social transitions by mitigating institutional free-riding. We can envision modelling organizational or financial reporting policies as a mechanism through which the observed fitness of a group can now explicitly depend on their institutional level ℓ . This signalling would inform groups not only about successful strategies but the required context for them to be successful (e.g. groups should only increase their institutional level in synchrony with local behaviour, so that institutional changes reliably generate group-level benefits). The model could be made more evolutionary as well. It could be implemented as a full multi-level selection model with emergent pressures of selection operating at both individual and group levels [26]. Or, future iterations could specify the social learning mechanisms in operation at each level.

Our model has the potential to shed light on a set of social processes that may be very influential in social evolution. If the dynamics that we demonstrate here are as prevalent as they appear, the implications of this study are deep and far-reaching. Moreover, our model provides important insights on the efficacy of approaches for addressing global challenges such as climate change. It provides a novel framework to study intervention regardless of the social scales at which actors operate.

Data accessibility. Model simulation code is available in C++ at <https://github.com/LaurentHebert/group-cultural-adaptation>, and has been archived within the Zenodo repository: <https://doi.org/10.5281/zenodo.5949710>.

Additional analyses are provided in electronic supplementary material [66].

Authors' contributions. L.H.-D. and T.M.W. developed the conceptual model, L.H.-D. wrote mathematical model and simulation, and analysed the model. G.S.-O. performed the perturbation analysis. T.M.W. wrote the paper. All authors contributed to the paper. N.J.G. and B.J.M. provided project funding and administration. Pre-print available on arXiv [67].

Competing interests. We declare we have no competing interests.

Funding. This research was supported by the US National Science Foundation (award EPS-2019470), Google Open Source under the Open-Source Complex Ecosystems And Networks (OCEAN) project, and the US Department of Agriculture (NIFA project no. ME022008).

Acknowledgements. The authors thank Adrian Bell, Jeremy Brooks and Vicken Hillis for inspiration.

References

- Bak-Coleman JB *et al.* 2021 Stewardship of global collective behavior. *Proc. Natl Acad. Sci. USA* **118**, e2114477118. (doi:10.1073/pnas.2114477118)
- Andreoni J, Nikiforakis N, Siegenthaler S. 2021 Predicting social tipping and norm change in controlled experiments. *Proc. Natl Acad. Sci. USA* **118**, e2014893118. (doi:10.1073/pnas.2014893118)
- Efferson C, Vogt S, Fehr E. 2020 The promise and the peril of using social influence to reverse harmful traditions. *Nat. Hum. Behav.* **4**, 55–68. (doi:10.1038/s41562-019-0768-2)
- Efferson C, Vogt S. 2018 Behavioural homogenization with spillovers in a normative domain. *Proc. R. Soc. B* **285**, 20180492. (doi:10.1098/rspb.2018.0492)
- Schmelz K. 2021 Enforcement may crowd out voluntary support for COVID-19 policies, especially where trust in government is weak and in a liberal society. *Proc. Natl Acad. Sci. USA* **118**, e2016385118. (doi:10.1073/pnas.2016385118)
- Sabatier PA, Weible CM. 2014 *Theories of the policy process*. New York, NY: Routledge.
- Shipan CR, Volden C. 2012 Policy diffusion: seven lessons for scholars and practitioners. *Public Adm. Rev.* **72**, 788–796. (doi:10.1111/j.1540-6210.2012.02610.x)
- Gilardi F. 2010 Who learns from what in policy diffusion processes? *Am. J. Pol. Sci.* **54**, 650–666. (doi:10.1111/j.1540-5907.2010.00452.x)
- Hodgson GM. 2002 The evolution of institutions: an agenda for future theoretical research. *Constit. Pol. Econ.* **13**, 111–127. (doi:10.1023/A:1015301101712)
- Muthukrishna M. 2020 Cultural evolutionary public policy. *Nat. Hum. Behav.* **4**, 12–13. (doi:10.1038/s41562-019-0780-6)
- Smaldino PE. 2014 The cultural evolution of emergent group-level traits. *Behav. Brain Sci.* **37**, 243–254. (doi:10.1017/S0140525X13001544)
- Bailey SA, Deneau MG, Jean L, Wiley CJ, Leung B, MacIsaac HJ. 2011 Evaluating efficacy of an environmental policy to prevent biological invasions. *Environ. Sci. Technol.* **45**, 2554–2561. (doi:10.1021/es102655j)
- McDonald P. 2006 Low fertility and the state: the efficacy of policy. *Popul. Dev. Rev.* **32**, 485–510. (doi:10.1111/j.1728-4457.2006.00134.x)
- Nelson TF *et al.* 2013 Efficacy and the strength of evidence of U.S. alcohol control policies. *Am. J. Prev. Med.* **45**, 19–28. (doi:10.1016/j.amepre.2013.03.008)
- Tews K, Busch P-O, Jörgens H. 2003 The diffusion of new environmental policy instruments. *Eur. J. Pol. Res.* **42**, 569–600. (doi:10.1111/1475-6765.00096)
- Lewis OA, Steinmo S. 2012 How institutions evolve: evolutionary theory and institutional change. *Polity* **44**, 314–339. (doi:10.1057/pol.2012.10)
- Richerson P, Boyd R. 2001 Institutional evolution in the holocene: the rise of complex societies. *Proc. Br. Acad.* **110**, 197–204.
- Bowles S, Choi J-K, Hopfensitz A. 2003 The co-evolution of individual behaviors and social institutions. *J. Theor. Biol.* **223**, 135–147. (doi:10.1016/S0022-5193(03)00060-2)
- Sunstein CR. 2014 *Why nudge? The politics of libertarian paternalism*. New Haven, CT: Yale University Press.
- Berger J. 2021 Social tipping interventions can promote the diffusion or decay of sustainable consumption norms in the field: evidence from a quasi-experimental intervention study. *Sustainability* **13**, 3529. (doi:10.3390/su13063529)
- Smaldino PE, Janssen MA, Hillis V, Bednar J. 2017 Adoption as a social marker: innovation diffusion with outgroup aversion. *J. Math. Sociol.* **41**, 26–45. (doi:10.1080/0022250X.2016.1250083)
- Schmelz K, Bowles S. 2021 How policies affect beliefs and preferences: the example of COVID-19 vaccination willingness. *VoxEU*. See <https://voxeu.org/article/how-policies-affect-beliefs-and-preferences>.
- Pellerano JA, Price MK, Puller SL, Sánchez GE. 2017 Do extrinsic incentives undermine social norms? Evidence from a field experiment in energy conservation. *Environ. Resour. Econ.* **67**, 413–428. (doi:10.1007/s10640-016-0094-3)
- Murayama K, Matsumoto M, Izuma K, Matsumoto K. 2010 Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proc. Natl Acad. Sci. USA* **107**, 20 911–20 916. (doi:10.1073/pnas.1013305107)
- Promberger M, Marteau TM. 2013 When do financial incentives reduce intrinsic motivation? Comparing behaviors studied in psychological and economic literatures. *Health Psychol.* **32**, 950–957. (doi:10.1037/a0032727)
- Waring TM, Goff SH, Smaldino PE. 2017 The coevolution of economic institutions and sustainable consumption via cultural group selection. *Ecol. Econ.* **131**, 524–532. (doi:10.1016/j.ecolecon.2016.09.022)
- Rogers EM. 2010 *Diffusion of innovations*. New York, NY: Simon and Schuster.
- Centola D. 2010 The spread of behavior in an online social network experiment. *Science* **329**, 1194–1197. (doi:10.1126/science.1185231)
- Centola D, Macy M. 2007 Complex contagions and the weakness of long ties. *Am. J. Sociol.* **113**, 702–734.
- Boyd R, Richerson PJ. 1985 *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.
- Beinhocker ED. 2006 *The origin of wealth: evolution, complexity, and the radical remaking of economics*. Boston, MA: Harvard Business Press.
- Boyd R. 2017 *A different kind of animal: how culture transformed our species*. Princeton, NJ: Princeton University Press.
- Henrich J. 2006 Cooperation, punishment, and the evolution of human institutions. *Science* **312**, 60–61. (doi:10.1126/science.1126398)

34. Bandura A. 1971 *Social learning theory*. New York, NY: General Learning Press.
35. Marceau V, Noël P-A, Hébert-Dufresne L, Allard A, Dubé LJ. 2011 Modeling the dynamical interaction between epidemics on overlay networks. *Phys. Rev. E* **84**, 026105. (doi:10.1103/PhysRevE.84.026105)
36. Hébert-Dufresne L, Mistry D, Althouse BM. 2020 Spread of infectious disease and social awareness as parasitic contagions on clustered networks. *Phys. Rev. Res.* **2**, 033306. (doi:10.1103/PhysRevResearch.2.033306)
37. Fischbacher U, Gächter S, Fehr E. 2001 Are people conditionally cooperative? Evidence from a public goods experiment. *Econ. Lett.* **71**, 397–404. (doi:10.1016/S0165-1765(01)00394-9)
38. Croson R, Fatas E, Neugebauer T. 2005 Reciprocity, matching and conditional cooperation in two public goods games. *Econ. Lett.* **87**, 95–101. (doi:10.1016/j.econlet.2004.10.007)
39. Baldassarri D, Grossman G. 2011 Centralized sanctioning and legitimate authority promote cooperation in humans. *Proc. Natl Acad. Sci. USA* **108**, 11 023–11 027. (doi:10.1073/pnas.1105456108)
40. Gächter S, Thöni C. 2005 Social learning and voluntary cooperation among like-minded people. *J. Eur. Econ. Assoc.* **3**, 303–314. (doi:10.1162/jeea.2005.3.2-3.303)
41. Molleman L, Quinones AE, Weissing FJ. 2013 Cultural evolution of cooperation: the interplay between forms of social learning and group selection. *Evol. Hum. Behav.* **34**, 342–349. (doi:10.1016/j.evolhumbehav.2013.06.001)
42. Bergin J, Bernhardt D. 2009 Cooperation through imitation. *Games Econ. Behav.* **67**, 376–388. (doi:10.1016/j.geb.2009.01.008)
43. Turchin P. 2016 *Ultrasociety: how 10,000 years of war made humans the greatest cooperators on earth*. Chaplin, CT: Beresta Books.
44. Richerson P *et al.* 2016 Cultural group selection plays an essential role in explaining human cooperation: a sketch of the evidence. *Behav. Brain Sci.* **39**, e30. (doi:10.1017/S0140525X1400106X)
45. van den Bergh JC, Gowdy JM. 2009 A group selection perspective on economic behavior, institutions and organizations. *J. Econ. Behav. Organ.* **72**, 1–20. (doi:10.1016/j.jebo.2009.04.017)
46. Francois P, Fujiwara T, van Ypersele T. 2018 The origins of human prosociality: cultural group selection in the workplace and the laboratory. *Sci. Adv.* **4**, eaat2201. (doi:10.1126/sciadv.aat2201)
47. Traulsen A, Nowak MA. 2006 Evolution of cooperation by multilevel selection. *Proc. Natl Acad. Sci. USA* **103**, 10 952–10 955. (doi:10.1073/pnas.0602530103)
48. Garcia J, van den Bergh CJM. 2011 Evolution of parochial altruism by multilevel selection. *Evol. Hum. Behav.* **32**, 277–287. (doi:10.1016/j.evolhumbehav.2010.07.007)
49. Simon B, Fletcher JA, Doebeli M. 2013 Towards a general theory of group selection. *Evolution* **67**, 1561–1572. (doi:10.1111/j.1558-5646.2012.01835.x)
50. Gavrillets S, Duwal Shrestha M. 2021 Evolving institutions for collective action by selective imitation and self-interested design. *Evol. Hum. Behav.* **42**, 1–11. (doi:10.1016/j.evolhumbehav.2020.05.007)
51. Ostrom E. 1990 *Governing the commons: the evolution of institutions for collective action*. Cambridge, UK: Cambridge University Press.
52. Wilson DS, Ostrom E, Cox ME. 2013 Generalizing the core design principles for the efficacy of groups. *J. Econ. Behav. Organ.* **90**, Supplement, S21–S32. (doi:10.1016/j.jebo.2012.12.010)
53. Shipan CR, Volden C. 2008 The mechanisms of policy diffusion. *Am. J. Pol. Sci.* **52**, 840–857. (doi:10.1111/j.1540-5907.2008.00346.x)
54. Bedson J *et al.* 2021 A review and agenda for integrated disease models including social and behavioural factors. *Nat. Hum. Behav.* **5**, 834–846. (doi:10.1038/s41562-021-01136-2)
55. Bettencourt LMA, Cintrón-Arias A, Kaiser DI, Castillo-Chávez C. 2006 The power of a good idea: quantitative modeling of the spread of ideas from epidemiological models. *Physica A* **364**, 513–536. (doi:10.1016/j.physa.2005.08.083)
56. Hébert-Dufresne L, Noël P-A, Marceau V, Allard A, Dubé LJ. 2010 Propagation dynamics on networks featuring complex topologies. *Phys. Rev. E* **82**, 036115. (doi:10.1103/PhysRevE.82.036115)
57. Renne J-P, Roussellet G, Schwenkler G. 2020 Two-thirds of COVID-19 deaths in the US could have been prevented through unified state and federal policies. *VoxEU*. See <https://voxeu.org/article/unified-state-and-federal-policies-could-have-prevented-two-thirds-covid-19-deaths-us>.
58. White ER, Hébert-Dufresne L. 2020 State-level variation of initial COVID-19 dynamics in the United States. *PLoS ONE* **15**, e0240648. (doi:10.1371/journal.pone.0240648)
59. Althouse BM, Wallace B, Case B, Scarpino SV, Allard A, Berdahl AM, White ER, Hébert-Dufresne L. 2020 The unintended consequences of inconsistent pandemic control policies. *MedRxiv*. (doi:10.1101/2020.08.21.20179473)
60. Hogler RL. 2015 *The end of American labor unions: the right-to-work movement and the erosion of collective bargaining*. ABC-CLIO. Google-Books-ID: WbgxBgAAQBAJ.
61. Sobel RS. 1995 Empirical evidence on the union free-rider problem: do right-to-work laws matter? *J. Labor Res.* **16**, 347–365. (doi:10.1007/BF02685762)
62. Hupper A. 2019 *Local food policy and consumer food cooperatives: evolutionary case studies*. PhD thesis, The University of Maine, Orono, ME.
63. Vreugdenhil H, Taljaard S, Slinger JH. 2012 Pilot projects and their diffusion: a case study of integrated coastal management in South Africa. *Int. J. Sustain. Dev.* **15**, 148–172. (doi:10.1504/IJSD.2012.044039)
64. Busch P-O, Jörgens H, Tews K. 2005 The global diffusion of regulatory instruments: the making of a new international environmental regime. *ANNALS Am. Acad. Pol. Soc. Sci.* **598**, 146–167. (doi:10.1177/0002716204272355)
65. Daley DM, Garand JC. 2005 Horizontal diffusion, vertical diffusion, and internal pressure in state environmental policymaking, 1989–1998. *Am. Pol. Res.* **33**, 615–644. (doi:10.1177/1532673X04273416)
66. Hébert-Dufresne L, Waring TM, St-Onge G, Niles MT, Corlew LK, Dube MP, Miller SJ, Gotelli NJ, McGill BJ. 2022 Source-sink behavioral dynamics limit institutional evolution in a group-structured society. Figshare.
67. Hébert-Dufresne L, Waring TM, St-Onge G, Niles MT, Corlew LK, Dube MP, Miller SJ, Gotelli N, McGill BJ. 2021 Source-sink cooperation dynamics constrain institutional evolution in a group-structured society. (doi:10.48550/arXiv.2109.08106)