Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-022-01095-7 OPEN

Collective narratives catalyse cooperation

Chaitanya S. Gokhale [1]✉, Joseph Bulbulia [2,3] & Marcus Frean [4]

Humans invest in fantastic stories—mythologies. Recent evolutionary theories suggest that cultural selection may favour moralising stories that motivate prosocial behaviours. A key challenge is to explain the emergence of mythologies that lack explicit moral exemplars or directives. Here, we resolve this puzzle with an evolutionary model in which arbitrary mythologies transform a collection of egoistic individuals into a cooperative. We show how these otherwise puzzling amoral, nonsensical, and fictional narratives act as exquisitely functional coordination devices and facilitate the emergence of trust and cooperativeness in both large and small populations. Especially, in small populations, reflecting earlier hunter-gatherers communities, relative to our contemporary community sizes, the model is robust to the cognitive costs in adopting fictions.

---

[1] Research Group for Theoretical Models of Eco-evolutionary Dynamics, Department of Evolutionary Theory, Max-Planck Institute for Evolutionary Biology, 24306 Plön, Germany. [2] School of Psychology, Faculty of Science, Victoria University of Wellington, Wellington, New Zealand. [3] Max Planck Institute for the Science of Human History, Jena, Germany. [4] School of Engineering and Computer Science, Victoria University of Wellington, Wellington, New Zealand. ✉email: gokhale@evolbio.mpg.de

Introduction

Human culture is a bewildering array of complex yet intuitively consistent, mutually beneficial arrangements. Humans are champions of cooperation, from the reproductive division of labour and alloparenting to collective foraging, information-sharing, education, and defence. Large-scale cooperation traces to a deep ancestral past, built on the back of highly social hunter-gatherer modes of existence (Bocquet-Appel, 2011; Bocquet-Appel and Bar-Yosef, 2008; Mcbrearty and Brooks, 2000). It has long been argued that a special class of narratives—moralising narratives—were a central feature of the Hominin cooperative adaptive complex. For example, according to the supernatural punishment hypothesis (Johnson, 2016; Swanson, 1960), natural and cultural selection targeted beliefs in the reality of supernatural policing, leading to a remarkable diversity of religious belief systems. According to these scholars, moral stories inculcate beliefs in supernatural enforcers, which in turn drives obedience to cooperative norms (Norenzayan et al., 2016; Shariff and Norenzayan, 2016). Despite the plausibility of adaptationist explanations for supernatural narratives, a key challenge remains explaining the cultural persistence of a-moral, morally ambiguous, and anti-social mythologies. Put differently, the class of narratives that humans transmit would appear to be larger than the class of stories that contain themes of supernatural enforcement. The emergence and conservation of non-moralising narratives, then, remains unexplained by moralising narrative theories.

An example of a narrative that would appear to lack explicit moral instructions is the myth about the primordial deities Rangi and Papa in Māori culture. According to this culturally conserved narrative, Rangi and Papa were once primordial beings who were connected in a loving embrace that encompassed all of reality. The world was created when one of their children, Tane, pushed his parents apart, thus giving rise to the separation of the sky (Rangi) and earth (Papa). Since then, according to this narrative the couple has lived in longing for each other with little or no regard for their children or mortals. It is difficult to derive any clear moral doctrine or instruction from this story. Indeed, many cultural narratives exhibit morally problematic features. For example, in Aztec mythology, the feathered serpent god Quetzalcoatl is described as having set himself on fire the day after an intoxicating evening in which he performed sexual acts on his sister Quetzalpetlatl. Both the god's incest and the self-immolation are arguably morally excessive, at least by standards to which mortals hold themselves accountable. Incest avoidance is a cultural universal, and does not appear to be challenging to motivate.

We might reject the theory that such narratives without moralising doctrines and supernatural policing agents lack moral effects. However, it is also possible that non-moral narratives might evolve to support prosociality in the absence of explicit moral instruction or moral enforcement. We assume the narratives are transmitted and conserved with fidelity across generations, and that this cultural regularity needs to be explained. Here, we offer a mathematical model that explains how arbitrary narratives may function as coordination devices facilitating the evolution of cooperation. Our model explains how folklore may evolve as a robust and reliable coordination device for cooperation, irrespective of the morality or sensibility of its content. This model is interesting because it parsimoniously explains how evolutionary dynamics could favour the otherwise puzzling evolutionary emergence and conservation of traditional narratives that lack explicit moralising instruction and enforcement.

Cooperation threatened by risk: The Stag Hunt. Imagine a group of hunters embarking on a stag hunt (Binmore, 1994; Calcott, 2008; Rousseau, 1755; Skyrms, 2003). Stag hunting yields a substantial payoff to all hunters, but the success depends on the collective action of all hunters. At any point, an individual hunger might decide to hunt a hare in place of the stag. Successful hare-hunting pays the hunter less than successful stag-hunting. However, hunting a hare is a safe bet. By contrast, stag-hunting is a safe bet only when all hunters choose to hunt stag. As such, stag hunting, though conditionally optimal, is threatened by risk-avoidance. In an all-or-none scenario, all hunters will fail if even one hunter seeks to avoid the risk that cooperation fails. This reasoning makes the default equilibrium of ubiquitous hare-hunting stable because propensities toward stag hunting face downward selective pressure. How can populations transit to cooperation (stag-hunting) under such circumstances?

We might look to the Theory of Mind, or the capacity to represent other people's mental states, for the solution. The consensus view is that the Theory of Mind was a critical step in achieving cooperation at scale (Tomasello and Carpenter, 2007). In a stag hunt, I can represent that it is in your interests to coordinate, and you can do the same. However, it has long been observed that the Theory of Mind may compromise cooperation in a coordination problem with risk. Merely knowing that risk-avoidance threatens the success of a stag-hunt might inspire risk-avoidance. For example, suppose individuals would choose stag with a probability of $p = 0.9$. Knowing this, agents with shared intentionality can reason that the chance of n choosing to hunt is $p^n \ll p$ and, being risk-averse, down-grade their own p accordingly. Theory of Mind allows partners to represent each other's uncertainty. The capacity to represent uncertainty in the setting of risk threatens coordination, even when all parties understand that coordination is in everyone's interests (Binmore, 2008; Binmore and Samuelson, 2001; Bulbulia, 2012; Bulbulia and Frean, 2010; Rubinstein, 1989). Examples of such coordination failures abound. During a pandemic, we understand that there is ample toilet paper for all. Nevertheless, we rush to the market, and a wrestling match ensues because we predict, correctly, that the pandemic will cue hoarding. Were we unable to represent the minds of others, no such problem would arise. Even successful collectives are fragile, and humans seem in need of reusable mechanisms to traverse from a world of risk aversion to one of cooperative alignment.

Distinct from the Theory of Mind—the capacity to represent the mental states of others—we define "inter-subjectivity" as the capacity of multiple individuals to share the same subjective beliefs (Scheff, 2006). Inter-subjective reality then is a belief, whether accurate or not, that is accepted by all parties concerned. Solving a coordination problem that involves risk, such as a stag hunt, requires an alignment of behaviours. We next introduce a model in which arbitrary fictions cause a form of inter-subjectivity that supports cooperative behaviours, even in the face of risks.

Model & results

Traversing the dilemma. A tribe of individuals decide to form a group to acquire food. The group consists of G individuals, each of whom could opt to pursue a hare or stag. The value of a hare is denoted by P_H and that of a stag by P_S. The probabilities of acquiring the prey are, however, contingent upon the group composition (see Supplementary Information). If enough people are hunting stag, the hunt will be successful and yield Π_S for each active participant. A failed hunt results in no payoff at all ($\Pi_S = 0$). All hare hunters have a fixed payoff of Π_H regardless of the group composition. The hare hunters payoff is, therefore, "risk-free", being achievable regardless of the decisions of others.

We label and track belief in two narratives, 1 and 2, respectively. Each individual has a preference for one or the other ($u^\star = 1$ or 2).

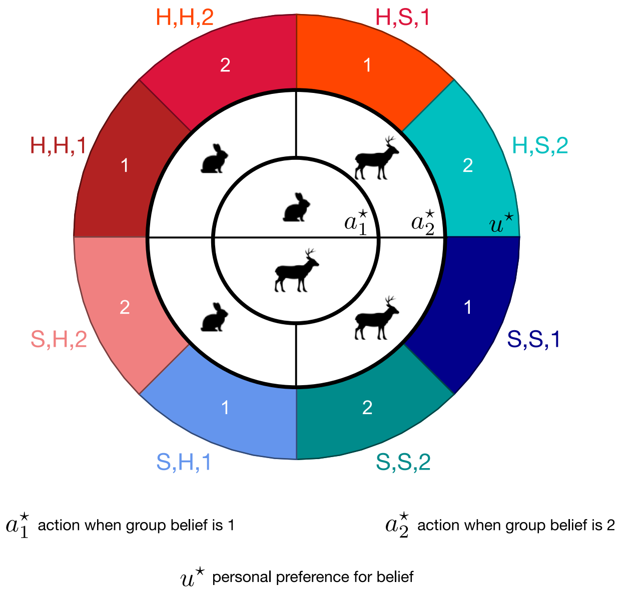


Fig. 1 The types of individuals in the world. Individuals are aware of the existence of narratives 1 and 2. They have a personal belief in either 1 or 2. The actions of the individuals could be *H* or *S* (shown as icons). Thus, the strategy of a focal individual $(\star)$ is defined by $(a_1^\star, a_2^\star, u^\star)$ where $a_i^\star$ is the hunting strategy of the focal individual when the group chooses to collectively believe in narrative *i* and $u^\star$ is the narrative that the individual personally believes in. Following their own narratives, the red-tinged strategies would hunt hares, and blue-tinged, stags.

Groups form, come to a consensus regarding their belief in narrative 1 vs. 2, and each member then adopts the behaviour (H or S) as per their actions in the chosen consensus belief. We denote $a_i^\star$ as the hunting action of the focal individual (H or S) when the group chooses to collectively believe in narrative *i*. The strategy of a focal individual $(\star)$ is then defined as a triplet given by $(a_1^\star, a_2^\star, u^\star)$. Thus, we have eight types of individuals to consider, as shown in Fig. 1. Given the consensus belief, the population has individuals who behave differently (H or S). Thus, while our model allows for conditional strategies, we do not restrict the strategies to follow only particular actions. Indeed conditional strategies would be the ones where a particular action is preferred over another, given a specific condition (in our case, the consensus belief), as is the typical use of conditional strategies (Gross and Repka, 1998). We begin in a world where everyone favours narrative 1 and is a hare hunter, hence a population consisting of type (H,H,1). Before setting off on a hunt, the group *G* arrives at a consensus about the narratives. A simple way to do this is to choose one member of the group and coalesce around their view - a dynamics of Groupthink (Bond, 2015). Humans frequently depend on the perceptions of others to inform themselves (Landauer, 1986). The notion of Groupthink captures the phenomenon in which a group minimises conflict by accepting a particular viewpoint without a critical evaluation of the decision or the alternatives. In this view, understanding is a contagion (Sloman and Fernbach, 2017) that allows for fast decision-making without conflicts. By selecting a random individual from the group, the narrative choice then amounts to a pure frequency-dependent ruling to reach a group decision.

Upon consensus, the individuals decide on their actions, given that the collective narrative is decided, and the consequences of the ensuing hunt unfold in the evolutionary dynamics. Thus, even individuals whose personal beliefs differ from the consensus can take action. We assume they go through a counterfactual “if – then” scenario and decide on an action. Such a rigorous reasoning mechanism is not a requirement of our model since we do not assume that the narrative conditions particular actions, all the possible actions are allowed to exist. Literature from psychology and previous models studying cooperation of evolution using counterfactual reasoning, however, provide support to such a mechanism (Catellani and Milesi, 2005; Pereira and Santos, 2019). To get an intuitive understanding of the evolutionary dynamics, it is instructive to first look at infinitely large populations, where the role of selection is highlighted (Apriasz et al., 2016). Although the payoff calculations are complicated (Supplementary Material), this approach allows us map out the directions of selection and the likely paths, which large finite populations might trace (Hauert et al., 2008). This approach, in turn, allows us to probe the more realistic case of smaller finite populations with insight (Traulsen et al., 2005, 2006).

Evolutionary dynamics in infinite populations. In infinitely large populations, deterministic dynamics reign supreme. In such a case, we calculate the average fitness of a population consisting only of one type of individual. Since there are eight different strategies, we calculate the fitnesses of all such homogeneous populations (Supplementary Material). We can perform a pairwise comparison of the fitnesses between the eight types to reveal full possible dynamics, assuming that mutation rates are low. The dynamics of the strategy evolution then proceeds in a seven-dimensional space whose vertices correspond to strategies, with intermediate points being a mixture of different strategies.

This simplex is projected in Fig. 2 as a two-dimensional graphic. The edges connect pairs of pure strategies with the arrows denoting the direction of selection when only those strategies are in play. Neutral dynamics (dashed connections) exist between strategies that choose the same narrative and the same action, such as (H,H,2) and (S,H,2), or even if they have different narratives but the same actions in both, such as (H,H,1) and (H,H,2). The neutrality emerges from the fact that the payoff eventually depends only on the actions and not the narrative that motivates it. Only two strategies ((S,H,1) and (H,S,2)) are dominant and selected over the corresponding Hare strategies ((S,H,2) and (H,S,1)). The population can thus move into the stag state where it can again evolve neutrally. The logic behind the two traversals beginning at (H,H,1) is as follows: (i) The bottom path allows for the neutral evolution of individuals that might hunt stags believing in narrative 2. This choice is not selected against because the consensus narrative is always 1, in which they hunt hares, and thus we have neutral drift (H,H,1) ⇄ (H,S,1). From (H,S,1) to (H,S,2) inter-subjectivity can break out en masse. If the group chooses narrative 2, the payoff is larger as they will get the stag. The belief in 2 increases until it becomes the inter-subjective reality – an accepted belief. (ii) The top path is possible in a similar fashion forming the stag-inducing inter-subjective reality of belief in narrative 1.

Following either path, we see that the population can go through a step in which the inter-subjectivity evolves (widespread belief in either narrative, the new one (bottom) or the pre-existing (top)) and catalyses a transition in behaviour. The dual transition is also evident in the future coexistence of the diversity of stag hunters, both believing in different narratives but still trusting each other to hunt the stag. Thus, while essential for mediating the transition, the exact choice of belief eventually becomes immaterial.

Alternative decision-making rules. Group formation itself can be a complex process. However, as this is not our focus, we assume it takes place with no assortment via random draws from the population. We have assumed that the hunting party decides

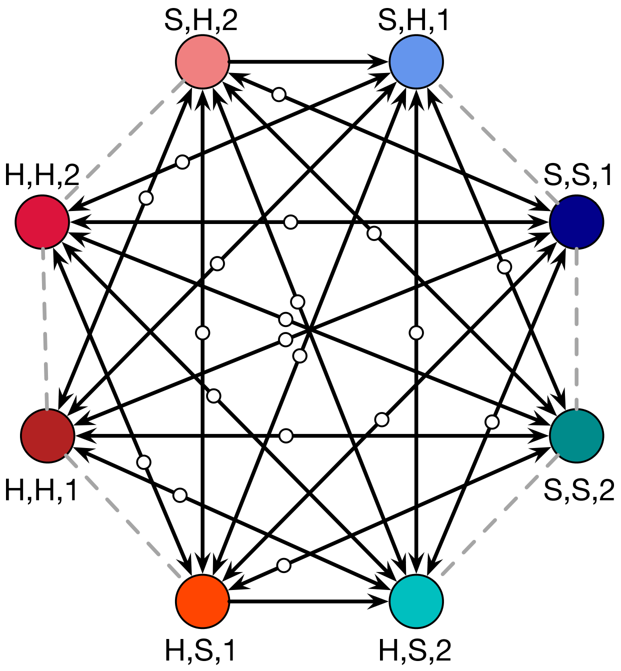
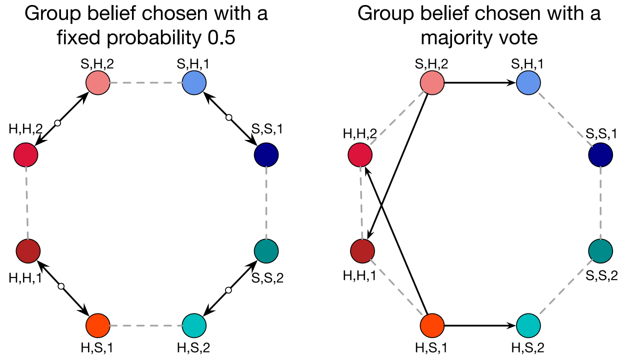


Fig. 2 Dynamics on the edges of the simplex. The dashed edges are neutrally stable whereas the directed edges show dominance. The edges with an open circle on them are all bi-stable (the circle marking the position of an unstable threshold as per Fig. SI.1), hence along those edges there would be no movement once the population is at a vertex. Once the population reaches the (H,S,2) or (S,H,1) state, it gets locked into the stag equilibrium and drifts between (H,S,2) - (S,S,2) - (S,S,1) - (S,H,1). Parameters are $G = 5$, $M = 4$, $P_S = 4$ and $P_H = 1$.


Group belief chosen with a fixed probability 0.5

Group belief chosen with a majority vote

Fig. 3 Dynamics on the edges of the simplex for alternative consensus making processes. Left. Instead of a frequency-dependent process as in Eq. (SI.1), if the consensus belief is decided with the flip of a fair coin then the eventual outcome is as shown in the figure. We do not show the links in the interior as all of them consist of an unstable equilibrium and are hence impassable. A population starting at (H,H,1) will be stuck in the hare equilibrium (H,H,⋆). The appearance of the unstable fixed points along the edges lock the population in pure hare or pure stag states. The consensus function $f(k, u^\star) = 0.5$. Qualitatively similar dynamics are obtained for other fixed values of $f(k, u^\star)$. Right. The consensus belief is chosen by a majority vote. The group belief is then the one which more than $G/2$ members prefer. In this case we see the appearance of a loop where (S,H,2)and (S,H,1)can lead back to the all hare hunters. However, in the long run, stag hunting in the ultimate sink for the dynamics. Now, the consensus function $f(k, u^{\text{star}})$ returns the chosen belief if the majority believes in it. For both the figures the other parameters are, $G = 5$, $M = 4$, $P_S = 4$ and $P_H = 1$.

on a narrative, and hence the ensuing actions of the involved individuals via a frequency dependent process (Eq. SI.1 in Supplementary Information (SI)). In this section describe the dynamics and the eventual distribution of the strategies when the group uses alternative methods for consensus forming. Charismatic leaders could impose their preference of narrative (Bulbulia and Frean, 2010), or the status-quo narrative might prevail, or everyone could decide for themselves. A simple majority vote is another appealing possibility supporting normative behaviour (Chalub et al., 2006; Ehrlich and Levin, 2005). The evolutionary dynamics of conformist behaviour can very well feed into the decision-making process (Denton et al., 2020). In Fig. 3 and in the Supplementary Material, we expound on two possible mechanisms of consensus reaching i.) Consensus is reached with a fixed probability of 0.5, i.e., irrespective of the group composition and ii) consensus is reached with a majority vote.

Fixed probability of realising a belief. If the group belief about the narrative choice is decided with a fixed probability, then the outcome of such a process is qualitatively different than the frequency dependent process as described earlier. For a fixed value of 0.5, i.e., at every consensus forming step the group belief is narrative 1 or 2 with probability 50%, we show the dynamics for infinitely large populations in Fig. 3 (left).

Majority vote. Whereas the frequency dependent nature of decision-making introduces a level of stochasticity in the system, a majority vote removes the ambiguity. If the number of individuals preferring narrative 1 form a majority, i.e., $k + \delta_{u^\star} > G/2$ then we assume that the group consensus is narrative 1 (else narrative 2).

Majority vote is denoted by a modification of the group decision function to:

$$f(k, u^\star) = \begin{cases} 1, & \text{if } k + \delta_{u^\star} > G/2 \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

For a given parameter set the resulting deterministic dynamics between the pure states are shown in Fig. 3 (right). Interestingly we see a feedback to the all hare states from the jump states of (H,S,1)and (S,H,2). In the long run, however, the population will escape this cycling and end up in the stag hunting state as seen in Fig. 3 (right).

The mechanism leading to the bypassing of the unstable equilibrium of a stag hunt game is similar to the scenario of the evolution of cooperation by tag-based selection (García et al., 2014). The narratives can be interpreted as different tags that the individuals bear and then can discriminate between partners—cooperators and defectors in a Prisoner's Dilemma. However, in our case, individuals do not discriminate based on the narratives they believe. Such strictly conditional strategies where the action set is restricted due to a condition may speed up the process. The disentangling of actions and narratives leads to highly complex dynamics, given the complicated nature of the payoff structure (see SI). The complexity prevents us from analysing the interior of the seven-dimensional simplex in detail. However, progress is possible using a more realistic picture of the model for finite populations, together with analytical tools and computer simulations.

Evolutionary dynamics in finite populations. Typical prehistoric communities were of small size as compared to our current community sizes. To analyse the dynamics of traits in tribes living together in countable population sizes needs different mathematics (Nowak et al., 2004). Let us constrain the size the population to N. Hence the number of individuals of a strategy are now integers, which can take values from zero to N.

We use a frequency-dependent Moran process to model the evolutionary dynamics. One of the crucial ingredients of the process is a strategy's average fitness. From a tribe of size N, a sample of G individuals decide to hunt. This sampling follows a hypergeometric distribution (Gokhale and Traulsen, 2010; Wu et al., 2013). Given the multiple conditionals in calculating payoffs and the population composition, the fitness calculation becomes a complicated matter of bookkeeping. We average over all possible types of groupings for calculating the average payoffs π_i for each strategy i (Appendix - Evolutionary dynamics in finite populations).

The average payoffs are converted to fitnesses as per $\psi_i = 1 + \omega\pi_i$. This conversion allows us to control the effect of the game on overall fitness. For $\omega = 0$, all strategies are equal with fitness of 1. If the interactions only have a minor effect on the overall fitness, then $\omega \rightarrow 0$. As the value of ω increases, the game has an increasing impact on fitness. Hence ω is known as the intensity of selection. The intensity of selection is a proxy for the relative importance of our complete model as compared to other activities that may impact the fitness of an individual (Venkateswaran and Gokhale, 2019). Thus, if hunting forms only a fraction of the total activities that define fitness, that fraction can be denoted by ω.

We choose an individual for reproduction proportional to the fitnesses and a random individual to die in each time step. To illustrate this step imagine a population consisting of only two strategies, i and j. The number of i individuals in the population is given by p; the number of j individuals is $N - p$. The probabilities that in a given time-step p decreases by one, stays the same, or increases by one, are given by,

$$
\begin{aligned}
T_{p,p-1} &= \frac{(1-\mu)\psi_j(N-p)+\mu\psi_i p}{\psi_i p+\psi_j(N-p)}\cdot\frac{p}{N}\\
T_{p,p+1} &= \frac{(1-\mu)\psi_i p+\mu\psi_j(N-p)}{\psi_i p+\psi_j(N-p)}\cdot\frac{N-p}{N}\\
T_{p,p} &= 1-T_{p,p-1}-T_{p,p+1}.
\end{aligned}
\tag{2}
$$

These transition probabilities also include the possibilities for mutations (μ) . For example $T_{p,p-1}$ can be read as, the probability that we select a strategy j individual proportional to fitness, which does not mutate or a strategy i individual proportional to fitness, which does mutate multiplied with the probability that a randomly chosen individual to die is of strategy i. This results in an overall increase in the number of j individuals by one. Considering only two strategies in a population we have these three possible outcomes. In a population where all strategies exist, there are therefore 56 possible transitions where the population composition can change (one strategy increases while another decreases) and the remainder probability where no change happens (same strategy increases and decreases). Following this algorithm of selective birth and random death, over time, the strategies with a larger fitness will spread and take over the population.

When an individual reproduces with a small probability (μ) the offspring might adopt a strategy different from its parent. If the new type has lower fitness than the resident, it will die out. However, a strategy with a larger average fitness than the resident has a chance of taking over the population. In a homogeneous population, a low mutation rate ensures that a mutant will have either gone extinct or taken over the population before a new mutant arises.

To demonstrate the dynamics of the process we consider a homogeneous population consisting of only (H,S,1) individuals. The Moran process, as above, will continue with replacing individuals without any discernible difference in the population composition. Now imagine if by chance a mutation occurs during reproduction leading to a (H,S,2) individual. The average fitness of an (H,S,2) individual is the same as that of a (H,S,1) individual in a population dominated by (H,S,1) . Thus, again the Moran process can spread the (H,S,2) mutation in a neutral manner (García et al., 2014). However, as (H,S,2) individuals increase in number, they will find themselves in groups dominated by other (H,S,2) individuals (Fig. 4). In such cases, the group consensus will be 2 and the action associated with it is to hunt stags. If the threshold number of hunters is met then the (H,S,2) individuals enjoy a larger payoff resulting in a larger average fitness. The Moran process will then select (H,S,2) individuals disproportionately more for reproduction, moving the population to a homogeneous (H,S,2) state. We explain this transition in Fig. 4.

Our model begins with a state in which all individual of a population of size N are (H,H,1) individuals. Mutations can lead to any of the strategies. For smaller population sizes, the stochastic effects are pronounced (population sizes explored in the Appendix). Given a population size, we can explore the effect of the different selection intensities on the evolutionary dynamics as well as the eventual abundances of the strategies as illustrated in Fig. 5.

Note that in the eventual distribution, mixtures of individuals believing in both realities are possible, indeed likely. Initially, lone strategists believing in a different narrative would be considered as irrational. However, since selection acts only via the action taken, the belief can spread neutrally as long as the action performed given the prevalent narrative remains hare hunting. Thus, although shared narratives facilitate the transition from hare hunters to stag hunters, the narratives themselves do not necessarily fix in the population but are merely a catalyst. Even non-believers in the narrative end up hunting stags. In this setting, to hunt stags becomes rational, irrespective of belief.

Costs. Notably, capacities to imagine and transmit shared narratives come with cognitive costs. In the Supplementary material, we systematically explore the case where only the new narrative believers pay the cognitive cost. To summarise that account, we find the production of narratives as a catalyst for coordination may evolve in the presence of cognitive costs. Notably, after enabling a transition to the stag equilibrium in finite populations, the belief in the new narrative disappears, leaving the population at the social optimum (Fig. SI.5).

Another aspect of the dynamics explored in smaller populations is the strategy update mechanism. Stemming from evolutionary arguments, we have used natural selection to drive the strategy dynamics rather than behavioural modifications. For example, when a group decides on a consensus narrative, the individuals who do not prefer the status quo narrative do not switch their preferences. Instead, they go with the consensus for the interaction with the relevant action. Natural selection ensures that strategies with a higher average payoff increase in frequency over time. An extension that incorporates active preference switching and the different methods of doing so would be interesting to explore (Traulsen et al., 2009) and bring our theory closer to behavioural studies. The sunk costs of a belief and the costs/benefits of relinquishing and adopting new beliefs would then need to be parameterised.

Furthermore, the finite population model also focuses on the magnitude of the benefit generated via collective action. In simpler evolutionary games, the selection intensity is used to measure the game's impact on the payoff of the strategies. When selection intensity is 0, all strategies are identical and reach equal abundance in finite populations. However, as we see in the case we present, this is not the same as equating the values of hares and stags. Crucially, we are interested in knowing how much more value a stag has than the hare? What is the impact of this differential on the evolutionary dynamics? A resolution is provided when we assess the fixation probabilities of the strategies in a pairwise manner (Nowak et al., 2004; Taylor et al., 2004). The payoff of the stag needs to be large enough for the whole endeavour to work, else the hare hunters reign supreme and no

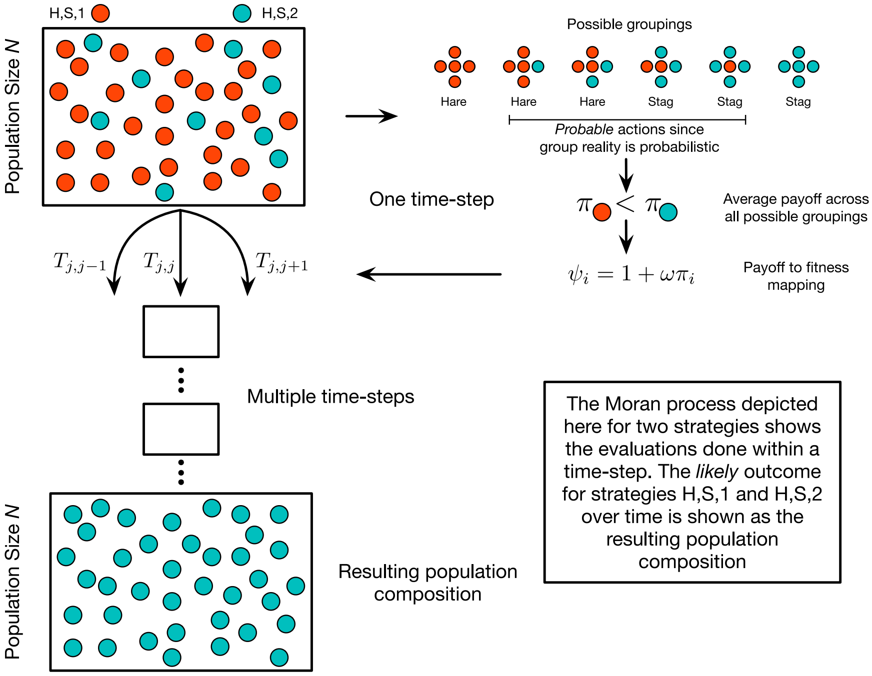


Fig. 4 Transitions in a finite population. In this figure, we highlight the transitions in a population of size N predominantly made up of (H,S,1) individuals and a small number of (H,S,2). For the sake of simplicity we have illustrated here an example with only two types. Mutations (here only between these two types) are incorporated in the transition probabilities. The groupings will typically consist of a majority of (H,S,1) resulting in narrative 1 as the group belief and hare hunting. The average payoff of a (H,S,2) individual is slightly larger than that of a (H,S,1) individual: although most groups will choose narrative 1, resulting in hare hunting, narrative 2 is occasionally chosen by those groups that contain (H,S,2) members, and such groups access the higher stag payoff. Over time this can result in (H,S,2) taking over the population. In general, multiple phenotypes can exist simultaneously following the same selection rule. In such cases there are fifty seven transition probabilities and mutations are possible between all the strategies. This can lead to a coexistence of multiple strategies as shown. Such eventual coexistence through a tumultuous transient phase can be seen in the Fig. 5.

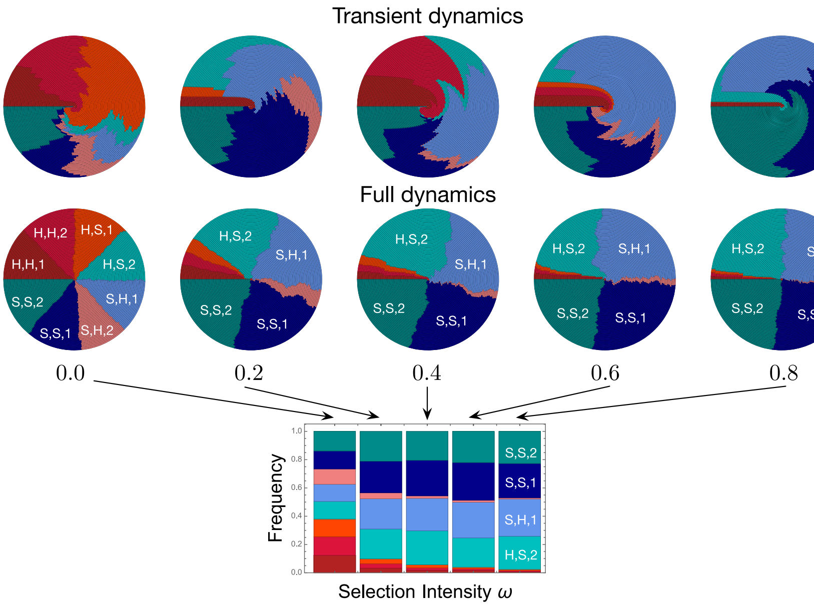


Fig. 5 Abundance plots for the eight strategies in a finite population. All simulations begin with a population consisting of (H,H,1). The circular figures are pie-charts evolving over time (initial condition inside and time progresses outwards). The rings of the charts show the population composition (denoted by the different colours) as it changes through time. *Transient:* For different selection intensities the initial conditions are all the same - starting in a (H,H,1) population (the centres of the circles). With a mutation rate of $\mu = 10^{-3}$ new strategies appear and spread in the population. Every 20th time-step up to 1000 time-steps are plotted, from inside to outside. *Full:* Over time (5×10^6 time-steps) the population reaches stationarity. Every 1000$^{\text{th}}$ time-step is plotted. The final distribution of strategies (outermost layer of the circles) is collated in the bar-chart. *Bottom:* The final distribution of the strategies at time-step 5×10^6 is plotted as a stacked bar chart for the different section intensities. When selection intensity $\omega = 0$ the dynamics is neutral and all the strategies evolve to similar abundances. When selection intensity is increased even slightly $\omega = 0.2$ the population is made up predominantly of stag hunters. Note that it is not necessary that all stag hunter share the same inter-subjective reality. Parameters are $N = 32$, $G = 5$, $M = 4$, $P_S = 4$ and $P_H = 1$.

benefits of collective action are gained. This despondent result is so because even if the stag and a hare have the same magnitude, a group of stag hunters can still lose the stag if not enough hunters are around. We show in the SI (Fig. SI.6) exactly how the magnitude of the payoffs changes the eventual outcomes in a finite population case.

Discussion

Summary. Previous research has found that moral mythologies can evolve to motivate pro-social behaviour (Chen, 2010; Henrich et al., 2010; Roes and Raymond, 2003; Shariff and Norenzayan, 2011, 2016; Sosis and Bulbulia, 2011). Numerous verbal or computational models exist, which examine the functional role of moral mythologies in the spread of cooperation (Roes and Raymond, 2003; Shariff and Norenzayan, 2011). Here, we present an mathematical model supported by stochastic simulations in finite populations that explicitly clarifies how any arbitrary narrative, whether explicitly moralising or not may nevertheless evolve to ensure efficient cooperation in risky coordination dilemmas at small and large social scales (Peoples HC, 2012). This model reveals an evolutionary pathway for utilising narratives that does not assume any moralising contents. As such, the model explains how otherwise puzzling non-moralising narratives may have evolved as exquisite adaptations to support human social life.

Caveats. First, although our focus has been to explain the evolutionary emergence and conservation of non-moral narratives, we do not rule out the prospect for cultural selection on explicitly moral narratives. Our analysis clarifies how non-moral narratives might evolve to support pro-social behaviours but does not claim whether moralistic contents give selective advantages.

Second, we do not attempt to explain the origin of basic capacities to imagine and transmit narratives: our account instead assumes such capacities (Cosmides and Tooby, 2000; Foster and Kokko; 2009; Guthrie et al., 1980). Indeed, our mathematical model assumes the actions and the ability for subjective inference are genetically encoded. In humans, evidence suggests that the potential for subjective reasoning might become regulated by genetic and cultural co-evolution deep in the hominin lineage, long before the emergence of agricultural and urban lifeways (Bulbulia, 2012; Foley and Lahr, 1997; Hill et al., 2009; Mcbrearty and Brooks, 2000; Powell et al., 2009; Sterelny, 2011).

Third, we have modelled the dynamics of the capacity to manifest narratives, assuming it as a single trait. However, human cognitive processing and behaviour are polygenic traits whose expression relies in part on locally varying cultural complexity (Norenzayan et al., 2016; Sosis and Alcorta, 2003). Linking cultural traits resulting in a cumulative cultural evolution and exploring the evolution of human cooperative 'packages' has been recently explored using agent-based simulations (Mesoudi and Thornton, 2018; Yeh et al., 2019).

Fourth, it is tempting to speculate about the role of narratives in our own time. Possibly, "post-truths", conspiracies, and a-moral ideologies might spread as mechanisms for coordinating individuals' behaviours into cooperative groups. However, whether modern narratives' popularity serves cooperative functions is not a question that we focus on (see Apriasz et al., 2016). Our model does not entitle us to speculate about the functions of stories in modern settings that involve multiple and overlapping coordination assurance mechanisms, which evolved from statecraft. Similarly, we resist speculating about the functions of traditional narratives in modern settings because our model does not clarify such questions. Whether stories that evolved to coordinate behaviours still do, is beyond the scope of the work presented here.

Future directions. We have assumed genetic traits for narratives. Looking ahead, future research should investigate gene-culture co-evolutionary dynamics. Speculatively, it might be possible to introduce a transmission of not only the strategies but the fictions themselves (Sosis and Bulbulia, 2011). The cultural evolution of successful stories, coupled with actions that provide a fitness advantage, could have led to thriving societies (Bietti et al., 2019). Furthermore, we have assumed that actions and preferences change together when mutations occur. Future work on decoupling the dynamics of action set from the belief would be an exciting avenue of research. This direction may help understand the causal relationships between actions and beliefs and the process of reinforcement of one or the other. Also, presently, we require that the same actions (H and S) are associated with the belief in the two narratives. We note that moving to truly conditional strategies; perhaps even the action space might yield different behaviours given a specific narrative belief (Gross and Repka, 1998; Repka and Gross, 1995).

Importance. Although the historical origins, and variation of human narrative thinking is not a topic for mathematical discovery alone, mathematical models such as ours are important because they reveal the limitations of intuitions for evolutionary discovery, allowing us to explore evolutionary obstacles and opportunities for social evolution. The purpose of this work has been to prove the concept that traditional cultures can evolve to assure cooperation in risky settings among unfamiliar partners (Bulbulia and Frean, 2010; Bulbulia and Sosis, 2011). Present theories either explain traditional mythologies as artefacts of unrelated cognitive features of the mind (Boyer and Parren, 2015), or require moralistic elements such as gods who punish anti-social behaviour (Beheim et al., 2019; Rockenbach and Milinski, 2006). Also noted are the traditional narratives that lack moralising elements typically highly motivating of cooperative behaviour (Atkinson et al., 2014). Our model offers a parsimonious account for the remarkable emergence and conservation of traditional stories as fortified mechanisms for the evolution of cooperation. This model is important because it reveals an evolutionary pathway for the conservation of traditional narratives as exquisitely adaptive functional adaptations for human social life.

Data availability

Appropriate computer code describing the model is available on Github https://github.com/tecoevo/beliefs.

Received: 29 April 2021; Accepted: 21 February 2022;

References

Apriasz R, Krueger T, Marcjasz G, Sznajd-Weron K (2016) The hunt opinion model-an agent based approach to recurring fashion cycles. PLoS ONE 11(11):e0166323

Atkinson QD, Latham AJ, Watts J (2014) Are big gods a big deal in the emergence of big groups? Relig Brain Behav 5(4):266–274

Beheim B, Atkinson QD, Bulbulia J, Gervais W, Gray RD, Henrich J, Lang M, Monroe MW, Muthukrishna M, Norenzayan A, Purzycki BG, Shariff A, Slingerland E, Spicer R, Willard AK (2021) Treatment of missing data determined conclusions regarding moralizing gods Nature 595:E29–E34

Bietti LM, Tilston O, Bangerter A (2019) Storytelling as adaptive collective sensemaking. Topic Cogn Sci 11(4):710–732

Binmore K (2008) Do conventions need to be common knowledge? Topoi 27(1-2):17

Binmore K, Samuelson L (2001) Coordinated action in the electronic mail game. Games Econ Behav 35(1-2):6–30

Binmore KG (1994) Playing fair: game theory and the social contract. MIT Press, Cambridge

Bocquet-Appel J-P (2011) When the world's population took off: the springboard of the neolithic demographic transition. Science 333(6):560–561

Bocquet-Appel J-P, Bar-Yosef O (2008) The Neolithic demographic transition and its consequences. Springer Science & Business Media

Bond M (2015) The power of others: peer pressure, groupthink, and how the people around us shape everything we do. Oneworld Publications

Boyer P, Parren N (2015) Threat-related information suggests competence: a possible factor in the spread of rumors. PLoS ONE 10(6):e0128421

Bulbulia J (2012) Spreading order: religion, cooperative niche construction, and risky coordination problems. Biol Philos 27(1):1–27

Bulbulia J, Frean M (2010) The evolution of charismatic cultures. Method Theor Stud Relig 22(4):254–271

Bulbulia J, Sosis R (2011) Signalling theory and the evolution of religious cooperation. Religion 41(3):363–388

Calcott B (2008) The other cooperation problem: generating benefit. Biol Philos 23(2):179–203

Catellani P, Milesi P (2005) When the social context frames the case. Counterfactuals in the courtroom. In: Mandel DR, Hilton DJ, Caterllani P (eds) The Psychology of Counterfactual Thinking. Routledge

Chalub FACC, Santos FC, Pacheco JM (2006) The evolution of norms. J Theor Biol 241:233–240

Chen DL (2010) Club goods and group identity: evidence from islamic resurgence during the Indonesian financial crisis. J Polit Econ 118(2):300–354

Cosmides L, Tooby J (2000) Metarepresentations: a multidisciplinary perspective. Ch. Consider the Source: The Evolution of Adaptationsfor Decoupling and Metarepresentation. Oxford University Press, New York

Denton KK, Ram Y, Liberman U, Feldman MW (2020) Cultural evolution of conformity and anticonformity. Proc Natl Acad Sci USA 117(24):13603–13614

Ehrlich PR, Levin SA (2005) The evolution of norms. PLoS ONE 3:943–948

Foley R, Lahr MM (1997) Mode 3 technologies and the evolution of modern humans. Cambridge Archaeol J 7(1):3–36

Foster KR, Kokko H (2009) The evolution of superstitious and superstition-like behaviour. Proc R Soc B: Biol Sci 276(1654):31–37

García J, van Veelen M, Traulsen A (2014) Evil green beards: tag recognition can also be used to withhold cooperation in structured populations. J Theor Biol 360:181–186

Gokhale CS, Traulsen A (2010) Evolutionary games in the multiverse. Proc Natl Acad Sci USA 107:5500–5504

Gross MR, Repka J (1998) Stability with inheritance in the conditional strategy. J Theor Biol 192(4):445–453

Guthrie S, Agassi J, Andriolo KR, Buchdahl D, Earhart HB, Greenberg M, Jarvie I, Saler B, Saliba J, Sharpe KJ et al. (1980) A cognitive theory of religion [and comments and reply]. Curr Anthropol 21(2):181–203

Hauert C, Traulsen A, Brandt H, Nowak MA, Sigmund K (2008) Public goods with punishment and abstaining in finite and infinite populations. Biol Theor 3:114–122

Henrich J, Ensminger J, McElreath R, Barr A, Barrett C, Bolyanatz A, Cardenas JC, Gurven M, Gwako E, Henrich N, Lesorogol C, Marlowe FW, Tracer D, Ziker J (2010) Markets, religion, community size, and the evolution of fairness and punishment. Science 327(Mar):1480–1484

Hill K, Barton M, Hurtado AM (2009) The emergence of human uniqueness: Characters underlying behavioral modernity. Evolution Anthropol: Issue News Rev 18(5):187–200

Johnson D (2016) God is watching you: how the fear of God makes us human. Oxford University Press, USA

Landauer TK (1986) How much do people remember? some estimates of the quantity of learned information in long-term memory. Cogn Sci 10(4):477–493

Mcbrearty S, Brooks AS (2000) The revolution that wasn't: a new interpretation of the origin of modern human behavior. J Human Evol 39(5):453–563

Mesoudi A, Thornton A (2018) What is cumulative cultural evolution? Proc Biol Sci 285(1880):20180712

Norenzayan A, Shariff AF, Gervais WM, Willard AK, McNamara RA, Slingerland E, Henrich J (2016) The cultural evolution of prosocial religions. Behav Brain Sci 39:e1

Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. Nature 428:646–650

Peoples HC MF (2012) Subsistence and the evolution of religion. Hum Nat 23(3):253–269

Pereira LM, Santos FC (2019) Counterfactual thinking in cooperation dynamics. In: Nepomuceno-Fernández Á, Magnani L, Salguero-Lamillar FJ, Barés-Gómez C, Fontaine M (eds.) Model-based reasoning in science and technology. Springer International Publishing, Cham. pp. 69–82

Powell A, Shennan S, Thomas MG (2009) Late pleistocene demography and the appearance of modern human behavior. Science 324(5932):1298–1301

Repka J, Gross MR (1995) The evolutionarily stable strategy under individual condition and tactic frequency. J Theor Biol 176(1):27–31

Rockenbach B, Milinski M (2006) The efficient interaction of indirect reciprocity and costly punishment. Nature 444:718–723

Roes FL, Raymond M (2003) Belief in moralizing gods. Evol Hum Behav 24(2):126–135

Rousseau J-J (1755) A discourse on inequality. Penguin Classics

Rubinstein A (1989) The electronic mail game: Strategic behavior under "almost common knowledge". Am Econ Rev 79(3): 385–391

Scheff TJ (2006) Goffman Unbound! (The Sociological Imagination). The Sociological Imagination. Routledge

Shariff AF, Norenzayan A (2011) Mean gods make good people: different views of god predict cheating behavior. Int J Psychol Relig 21(2):85–96

Shariff AF, Norenzayan A (2016) God is watching you. Psychol Sci 18(9):803–809

Skyrms B (2003) The stag-hunt game and the evolution of social structure. Cambridge University Press, Cambridge

Sloman S, Fernbach P (2017) The knowledge illusion: why we never think alone. Riverhead Books

Sosis R, Alcorta C (2003) Signaling, solidarity, and the sacred: the evolution of religious behavior. Evolution Anthropol Issue News Rev 12(6):264–274

Sosis R, Bulbulia J (2011) The behavioral ecology of religion: the benefits and costs of one evolutionary approach. Religion 41(3):341–362

Sterelny K (2011) From hominins to humans: how sapiens became behaviourally modern. Philos Trans R Soc London. Series B, Biol Sci 366(1566):809–822

Swanson GE (1960) The birth of the gods: The origin of primitive beliefs. Vol. 93. University of Michigan Press

Taylor C, Fudenberg D, Sasaki A, Nowak MA (2004) Evolutionary game dynamics in finite populations. Bull Math Biol 66:1621–1644

Tomasello M, Carpenter M (2007) Shared intentionality. Dev Sci 10(1):121–125

Traulsen A, Claussen JC, Hauert C (2005) Coevolutionary dynamics: From finite to infinite populations. Phys Rev Lett 95:238701

Traulsen A, Claussen JC, Hauert C (2006) Coevolutionary dynamics in large, but finite populations. Phys Rev E 74:011901

Traulsen A, Hauert C, De Silva H, Nowak MA, Sigmund K (2009) Exploration dynamics in evolutionary games. Proc Natl Acad Sci USA 106:709–712

Venkateswaran VR, Gokhale CS (2019) Evolutionary dynamics of complex multiple games. Proc R Soc B: Biol Sci 286(1905):20190900

Wu B, Traulsen A, Gokhale CS (2013) Dynamic properties of evolutionary multiplayer games in finite populations. Games 4(2):182–199

Yeh DJ, Fogarty L, Kandler A (2019) Cultural linkage: the influence of package transmission on cultural dynamics. Proc Biol Sci/R Soc 286(1916):20191951

Acknowledgements

We acknowledge Małgorzata Fic for a careful reading of a previous draft. Funding to J.B. from the Templeton Religion Trust 0196 is acknowledged. Funding to C.S.G. and the sabbatical support to M.F. from the Max Planck Society is graciously acknowledged.

Author contributions

C.S.G. and M.F. derived the mathematical and computational model. All author analysed the results and wrote the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The author(s) declare no competing interests.

Ethical approval

Not applicable.

Informed consent

Not applicable.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-022-01095-7.

Correspondence and requests for materials should be addressed to Chaitanya S. Gokhale.