Evolving the Core Design Principles: The Coevolution of Institutions and Sustainable Practices

Timothy M. Waring[1], Sandra Goff[1], and Paul E. Smaldino[2]

[1]University of Maine, Orono, ME USA
[2]Johns Hopkins University, Baltimore, MD USA
timothy.waring@maine.edu,sandra.goff@maine.edu,paul.smaldino@gmail.com

Abstract. The sustainable use of limited resources presents a looming cooperation problem for the global human population. A number of researchers have proposed that certain social norms and institutions can stabilize resource conservation behavior within groups, but a mechanism by which those norms might emerge is as yet unestablished. We speculate that such norms may emerge via a process of cultural evolution amongst social-ecological systems. We present a model of endogenous group formation to assess whether cultural multilevel selection may facilitate the emergence of sustainable cultures, even when individual conservation behavior remains costly. We show that conservation and population persistence is enhanced by the endogenous emergence of two social norms: boundary defense (or 'property') and cooperative production, and that the associated evolutionary process requires selection at the group level. We also show that longer-lasting institutions rely on the ability of individuals to form clearly-marked social groups. Our findings provide a new dynamic theory of the emergence of sustainable institutions and societies, and also have broad impacts for theories of human social evolution.

Keywords: sustainability, group selection, cultural evolution, Ostrom

1 Introduction

We live in a world in which the likelihood of environmental catastrophe looms larger every day. We are pulling food, fuel, and forest out of the earth at unsustainable rates. Industrial fisheries, for instance, have reduced the biomass of large predatory fish in the North Atlantic to only 10% of pre-industrial levels [1]. Our global climate has just entered a state that it has not seen for many thousands of years. It is therefore vital to understand how to advance the widespread adoption of sustainable practices. A persistent problem when considering the stewardship of common resources, however, is the vulnerability of socio-ecological systems to free riders. Although countless programs promote sustainable or low-impact living, any proposal for a behavioral or institutional solution is useless unless one can show that (1) we can get there from here (i.e., the solution can evolve from the present conditions), and (2) we can stay there when we do (i.e., the solution is resistant to invasion by free-riding strategies).

Using an agent-based model, we investigate the co-evolution of sustainable practices with normative institutions that allow groups to maintain such practices. Our approach relies on the theory of cultural evolution [2, 3], which accounts for the influences of culture and social learning on human behavior and organization. This is based on the recognition that natural selection does not require genes, but only a population of individuals (or other units of selection) among which there is variation, transmission of traits between individuals, and a selection for certain traits. Before introducing the model, we will first discuss institutions related to sustainability, followed by an introduction to key evolutionary principles.

1.1 Institutions for sustainability: Ostroms core design principles

Institutions are the formal and informal rules that govern human social behavior [4]. The primacy of institutions in determining the success of resource management regimes is well documented. Elinor Ostrom developed a set of institutional 'design features' that were commonly associated with successful resource management [5, 6]. These eight principles, abbreviated in Table 1, emerged from decades of comparative case study research. Groups that successfully managed common pool resources without succumbing to the "tragedy of the commons" [7] appeared to follow some subset of these principles, each of which can be characterized as an institution or institutional feature.

Although these institutions seem to be associated with effective sustainable management of common resources, it is unclear how they would arise initially. Many normative institutions and traditions themselves require cooperation and coordination among group members. Cooperative endeavors are often vulnerable to free riders who might cause such endeavors to fail before institutions that safeguard cooperation have the chance to emerge. It is therefore important to adopt a perspective that accounts for processes of social change in order to understand how institutions like Ostrom's design principles could have emerged, and perhaps even to understand how they could spread in a modern world so in need of sustainable practices.

1.2 Evolutionary principles

To guide the understanding and implementation of sustainable practices, we need to understand how they can arise and be maintained. A cultural evolutionary approach is useful here, because it is based on understanding the dynamics of social norms and cultural institutions. One important concept is the *evolutionarily stable strategy* (ESS) [8]. In order for a phenotypic trait – or "strategy" – to be evolutionarily viable, it must be able to resist an invasion by individuals or groups with competing strategies. Importantly, whether or not a strategy is an ESS depends not only on the pool of competing strategies, but also on existing social institutions and socio-spatial structure. For example, if a population is

Table 1. Ostrom's eight core design principles (adapted from [5]).

	Principle	*Description*
1	Boundaries	Clearly defined boundaries (effective exclusion of external un-entitled parties);
2	Rules	Rules regarding the appropriation and provision of common resources that are adapted to local conditions;
3	Collective-choice	Collective-choice arrangements that allow most resource appropriators to participate in the decision-making process;
4	Monitoring	Effective monitoring by those who are part of or accountable to the appropriators;
5	Sanctions	Graduated sanctions for resource appropriators who violate community rules;
6	Conflict Resolution	Mechanisms of conflict resolution that are cheap and of easy access;
7	Self-determination	Self-determination of the community recognized by higher-level authorities;
8	Nested Organization	In the case of larger resources, organization in the form of multiple layers of nested enterprises.

well-mixed and interactions occur at random, then a free-riding strategy cannot be invaded by a purely cooperative strategy. However, even without explicit institutions of monitoring or punishment, cooperation can arise if there is rigid spatial or group structure [9–11]. The presence of punishment or reputational institutions can also stabilize cooperation [12–14].

Importantly, selection pressures for cooperation typically vary by the level at which selection operates. Well known psychological and cultural forces stabilize cultural groups, minimizing within-group differences and accentuating differences among groups, which allows selection to act at the group level [15–17]. Group selection can thereby select for cultural traits that would be maladaptive individually but can be favored by selection because they produce a group benefit [18]. Recently, *cultural group selection* has emerged as a general framework to explain the group-driven evolutionary processes that influence cultural and institutional evolution [15, 19, 20]. Henrich [15] lists four mechanisms that drive cultural group selection; differential imitation, differential migration, differential reproduction, and differential extinction. In all cases, the strength of group-level selection increases as fitness relevant between-group differences increase and as variation among individuals within a group declines [19].

Research on cultural group selection has largely been in the context of basic science research on the evolution of human institutions and propensities (e.g., [21–23]). However, there is scope for using this framework in more applied settings, such as the evolution of resource management institutions.

1.3 The coevolution of institutions and sustainable practices

We consider the coevolution of three cultural traits that interact with individual-level resource usage. The behavior we are most interested in is whether or not an individually costly practice of sustainable harvesting can be maintained in a population via the mechanisms of cultural group selection operating on relevant cultural traits. The traits we investigate are:

1. Shared production, whereby individuals share resources with others to produce a surplus;
2. Boundary defense, whereby individuals pay a cost to defend their resources from others;
3. Group markers, whereby individuals of different groups can be identified as such, allowing differential action towards in- and out-group members.

Shared production In an uncertain world in which individuals may experience droughts and gluts of resources, an institution through which current resources are shared and create surplus resources can mitigate this risk [24, 21]. We allow a simple cooperative economic production institution to evolve endogenously among individuals. Cooperative surplus production occurs through a two-person public goods game, so that there is no guarantee of even benefits.

Boundary defense The ability to declare a resource ones own, and to defend it from the hands of others, is common to most societies. Rousseau [25] wrote "The first man who, having enclosed a piece of ground, bethought himself of saying 'This is mine,' and found people simple enough to believe him, was the real founder of civil society." The first of Ostrom's [5] design principles is the existence and maintenance of clearly defined resource use boundaries (see Table 1). We allow the institution of boundary defense (or property) to evolve endogenously among individuals, whereby individuals can pay a cost to prevent others from exploiting resources on their land.

Shared production and boundary defense are costly and cooperative behaviors. Shared production is modeled as a symmetrical, dyadic public goods game in which individuals stand to benefit from a value-added collaborative process, but their contributions to the shared production may be taken advantage of by partners who do not contribute. Maintaining property is individually costly, and is also a cooperative act when property is shared with group members, because there is no guarantee that group members will pay the same costs or reciprocate by allowing harvesting on their patch.

Group markers Ethnic groups and other cultural divisions are often explicitly marked, and individuals identify themselves outwardly as a member of their group. One theory of human evolution posits that humans evolved the ability to form well-marked ethnic or social groups because doing so facilitates solving collective action problems such as resource procurement and inter-group conflict. Group markers can allow selective cooperation within groups [26], or facilitate increased payoffs due to the greater ease with which individuals with common knowledge can coordinate [27]. Ostrom also discusses her principle of clear boundaries such that the identity of the group and the boundaries of the shared resources are clearly delineated ([6], p. 2). This implies that the ability to easily distinguish in- and out-group members is an important component of boundary defense. By facilitating positive assortment among members of the same cultural groups, group markers promote decreased within-group variation and increased between-group variation, which strengthens selection at the level of groups. We investigate the inclusion of group markers, which allow individuals to preferentially share with members of their own group ? with whom they are also likely to share behavioral characteristics and to preferentially defend their resources only against out-group members.

1.4 Modeling the coevolution of institutions and sustainable practices

We apply the framework of cultural group selection to the emergence and persistence of sustainable practices, with a focus on the endogenous coevolution of Ostrom's first design principle: boundaries. We present an agent-based model in which individuals must harvest limited but renewable resources in order to thrive and survive. Sustainably-harvesting individuals avoid over-exploiting local resources, but can be out-competed in the short run by exploitative over-harvesters. We consider the simultaneous coevolution of social institutions and group markers.

Other multilevel selection models of social-ecological systems evolution have in general not allowed for endogenous group formation or allowed groups to vary in size [21, 28–31]. A recent model by Smaldino et al. [32] included endogenous group formation and variable population size, but did not investigate the influence of institutions. In our model, groups emerge endogenously from spatial organization. We do not fix population or group size, spatial extent, or group membership; therefore group selection arises entirely through individual-level processes.

We render three hypotheses:

1. When conservation is individually costly, population persistence and sustainable resource management requires social norms or institutional features, including shared production and boundary defense.
2. When resource conservation, sharing, and boundary defense are individually costly, they should not be stable without selection acting at the level of the social group.
3. For group-level selection to occur, social group markers must be present to promote preferential behavior toward group members.

2 Methods

The model simulates a simple society in which individuals harvest resources in order to survive and reproduce. Interactions occur on a discrete square lattice with toroidal topology. The lattice is comprised of patches, each of which grows a renewable resource over time according to a logistic function with a maximum growth rate (r) and a carrying capacity (K). Agents occupy only one patch at a time and, similarly, patches may have only a single occupant within a given time step. Agents can harvest resources from the patch they occupy as well as the eight surrounding patches (their Moore neighborhood). Agents have a binary harvesting trait (h_i) which may take either high or low states. An agent with a high harvest trait (H_H) prefers to harvest more than the maximum sustainable yield (MSY) for a single patch, while an agent with a low harvest trait (H_L) prefers to harvest less than the MSY, and is by definition sustainable.

Agents interact through the two additional behavioral traits, sharing (s_i) and defense (d_i), and a social marker trait (g_i). The behavioral traits direct agents to share access to the resources on their patch with, or defend them against, one of three sets of neighbors: all neighbors, ingroup neighbors only, or no one. The marker trait (g_i) may take one of nine possible values. The agent traits s_i, d_i, h_i, and g_i are allowed to change both within an agents lifetime through payoff-biased local imitation, as well as between generations through mutation of traits of offspring. These behavioral traits represent social norms that affect the actions taken by agents.

The model is initialized with nine groups of twelve identical, spatially clustered agents to allow for all combinations of the two ternary behavioral traits (sharing, defense). Agents in each group are assigned a combination of the two behavioral traits, and groups may additionally be assigned a unique marker. All agents are initialized with low harvesting preferences (H_L), and are assigned an initial random amount of harvested resource. Patches are initialized with a random amount of resource. In treatments where markers are active, markers are by default correlated with behavioral trait combinations (at least initially).

2.1 Simulation process order each time step

Each time step in a simulation run consists of the following processes, executed in sequential order.

1. *Patch defense*: Agents pay a cost to defend their focal patch from neighbors, determined by the agents exclusion criteria (no one, non-group members, everyone).
2. *Harvesting*: Agents harvest raw resources from their focal and neighboring patches according to their harvesting preference.
3. *Sharing*: Agents have the opportunity to collaborate with a neighbor to produce value-added processed resources.
4. *Pay cost of living*: Agents lose resources in order to stay alive.

5. *Death*: Agents die if their resources drop below zero or due to random events, the probability of which increases with age.
6. *Reproduction*: Agents with sufficient resources reproduce if there is an open patch in their Moore neighborhood.
7. *Imitation*: Agents may imitate the traits of local others according to a wealth bias.
8. *Patch growth*: Patch resources increase in a logistic fashion if they are below maximum.
9. *Cap resources*: Agent resources are reduced in compliance with a cap on accumulated resources.

At the beginning of each time step agents choose whether or not to defend the resources on their patch from neighboring harvesters. Defense is important because although agents are only allowed to occupy a single patch, they are able to harvest resources from any or all of the patches in their Moore neighborhood (the group of 8 patches directly surrounding their own patch), their commons. The defense trait determines whether an agent defends against no one (N), out-group members (O), or all individuals (A). Agents pay a fixed cost for each neighbor against whom they defend their property. Once all agents have chosen their defense strategy, agents harvest according to their harvesting norm (H_H or H_L).

Individuals can store only a limited amount of harvested resources according to the cap, R_{RL}, which implements the assumption that raw resources are perishable. Once all of the agents have had a chance to harvest, individuals may then have the opportunity to engage in a cooperative production activity that converts harvested resources into processed resources. The production process creates a surplus so that the resulting processes resources exceed the value of the initially supplied harvested resources. However, the process of production is a cooperative activity requiring two participants. We model this as a public goods game in which both contributions are summed, multiplied by a factor of production, Θ, and divided evenly among the two neighbors. To create a social dilemma the factor of production must be set so that $1 < \Theta < 2$. Here we set Θ to 1.5. To meet the costs of living and reproduction, individuals use processed resources first and harvested resources only when their processed resources are insufficient. This is because only the harvested resources may be used as contributions to the public goods game. An individual's behavior in the public goods game is determined by their sharing norm, s_i, which interacts with their group marker to determine the set of neighbors with whom they may possibly join in cooperative production. Sharing norms determine partner selection for production and may be either no one, everyone, or group members only.

Next, a fixed cost of living (C_L) is deducted from each agent's resource supply, as in [32]. Agents who are not able to pay the full cost die and are removed from the simulation. Agents can also die via a stochastic age-based function in which the probability of dying increases with age (see ODD documentation). Surviving agents reproduce asexually if they have resources in excess of the cost of reproduction (C_R) plus a parental survival margin and if there is an empty patch within their Moore neighborhood. Offspring inherit parental behavioral traits and markers, each of which mutates with probability μ.

Finally, each individual attempted, with probability λ, to learn from the wealthiest peer agent within a local radius, regardless of the traits or marker status of the model agent. Depending upon the treatment conditions being tested, each agent alters its harvesting, sharing, property, and/or marker traits to match those of the agent within a globally specified radius of the focal agent's patch with the highest total accumulated resources (raw and processed).

This sequence of agent-focused events is followed by the regrowth of patch resources based upon a logistic growth function such that each patch has a maximum carrying capacity of raw resources. See the ODD documentation for further model specification.

2.2 Design of simulation experiments

We conducted simulation experiments to address our three hypotheses. Five simulation treatments varied the existence of social markers or the availability of behavioral or institutional options for either shared production or property defense. Treatments were constructed according to Table 2. The global parameters used within these treatments can be found in the ODD appendix.

Table 2. Simulation treatment conditions and hypothesis tests.

#	**Treatment**	**Production**	**Property**	**Markers**	**Hypothesis**
1	no norms			Y	1
2	no property	Y		Y	1
3	no production		Y	Y	1
4	both norms	Y	Y	Y	1, 2, 3
5	no markers	Y	Y		3

For each treatment within each hypothesis test we ran 500 simulations for 1000 time steps under varying conditions, with benchmark parameter values. All runs were initialized with nine groups of 12 individuals each Individuals were homogenous within groups, and all individuals are set to harvest sustainably. High harvesting behavior was allowed to invade through both genetic and cultural transmission, and always did. Parameter values used in our simulations are listed in the ODD Appendix.

3 Results

In all simulations, populations underwent an initial boom as they consumed all available resources and grew to fill the entire spatial extent, followed by a crash when those resources became depleted (Fig. 1). As expected, once resources were depleted and populations crashed, the fitness advantage to harvesting at a lower rate aided the spread of conservation behavior (low harvesting) in the populations that manage to survive the initial crash (Fig. 2).

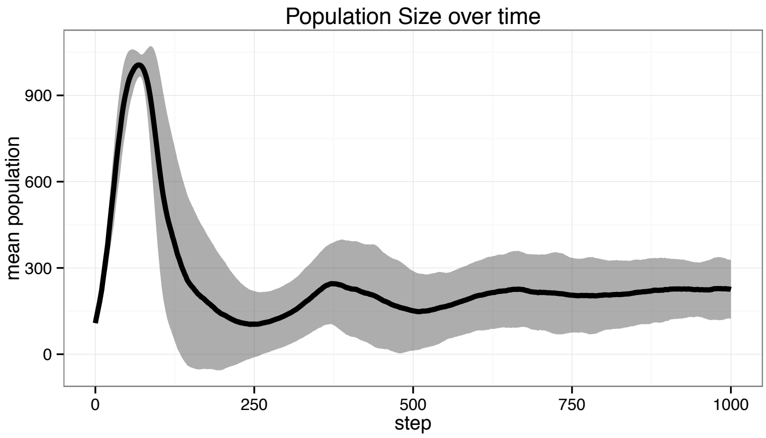

Fig. 1. Simulated populations undergo an initial boom and bust in all conditions. (1000 simulation mean ± SD)

3.1 Hypothesis 1: Long term persistence requires property and production norms

Over the long run, our model requires that agents consume resources sustainably for population persistence, but does not guarantee the emergence of this behavior. Here we ask if property and cooperation norms are required to stabilize altruistic resource conservation practices (H1). To evaluate this hypothesis we compared treatments in which social norms were allowed to evolve endogenously with ones in which they were fixed at a base institutional level. Treatments 1-4 (Table 2) examine each of the four possibilities for the norms of cooperative production and boundary defense. These social norms did prove necessary for population persistence and the stabilization of conservation behavior. When neither production nor boundary defense norms were allowed to evolve endogenously, all populations went extinct by ∼260 time steps. When only cooperative production could evolve, all populations went extinct by ∼700 time steps. When only property rights norms evolved, ∼12.5% of populations persisted for the entire 1000 time steps, and when both property and production norms were allowed to vary endogenously, ∼40% of populations persist through the entire time course (Fig. 3).

Persistence in the model was determined by the frequency of sustainable (low) harvesters in the population over time. Examination of the model dy-

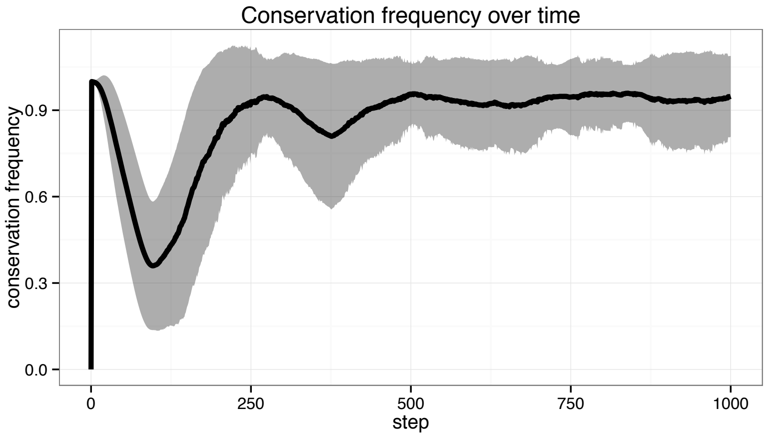

Fig. 2. Conservation frequency grows to high levels in populations that persist. (1000 simulation mean ± SD)

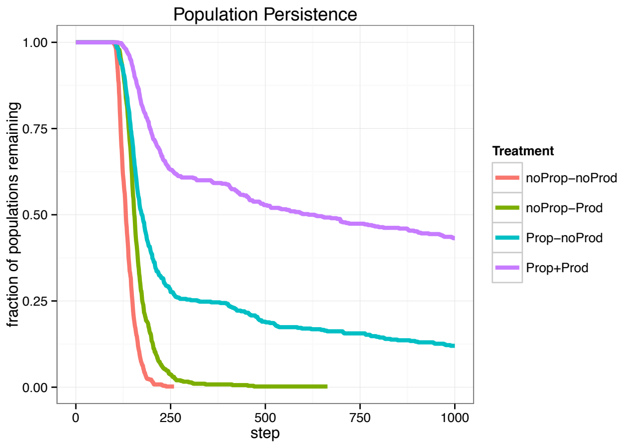

Fig. 3. Population survivorship curves decline for all treatments, but populations persist when the norms of property and production are allowed to evolve.

namics indicates that initial conservation behavior is quickly invaded by high consumption behavior in all treatment scenarios. Thereafter, however, the the capacity for social norms of production and boundary defense could, in some instances, stabilize conservation and thereby potentiate population persistence in the long run.

The fraction of populations surviving declined over time for all treatment conditions. This pattern suggests norms of boundary defense and cooperative production are not sufficient to stabilize sustainable practices in the long run. Other research on the evolution of cooperative institutions suggests that additional mechanisms such as conformity and punishment [12] might further stabilize cooperation and population persistence. One intriguing possibility is that the additional institutional traits and social norms identified by Ostrom [5] may also slow this rate of decline by stabilizing resource management. Future simulation studies will include these additional variables.

Examining the relationship between conservation frequency and population survival over time, we find that while all populations experienced an initial drop in conservation, those without necessary social norms could not bounce back by evolving a resurgence in conservation behavior, while those treatments with endogenous norm evolution more frequently rebound, and as a result, may persist (Fig. 4). Thus, we find that property rights are sufficient for long-term population persistence (1000 time steps) but that when production and property norms both evolve, chances of populations survival increases three-fold.

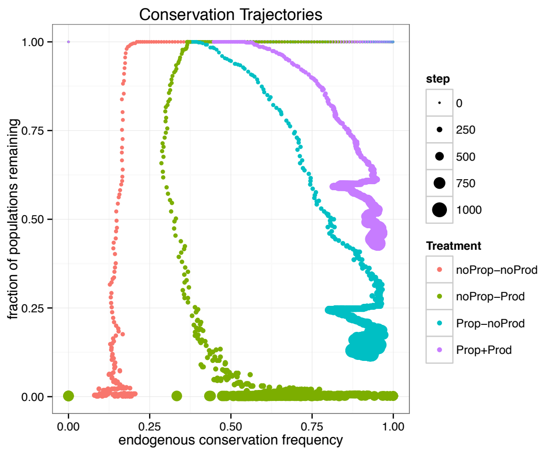

Fig. 4. Population survivorship as a function of endogenously evolved conservation behavior, by time and treatment.

3.2 Hypothesis 2: Group selection stabilizes property and cooperation norms

Having determined that the endogenous evolution of property and cooperation norms is necessary to stabilize altruistic resource conservation practices and maintain populations (H1), we now ask if the evolution of those norms is due to selection between endogenously forming groups or amongst individuals (H2). To test this question we conducted two treatments; one in which the initial groups are given distinct social markers that are inherited (and mutate at a low rate), and a second in which the initial groups are given a single marker and markers are not allowed to evolve. With distinguishing groups markers, ∼40% of populations persisted, compared with only ∼10% without group markers.

The Price equation [33, 34] allows us to separate the components of selection into those acting at the level of the individual and those acting at the level of the group. We computed Price's components of selection for the trait of conservation behavior (low harvesting), which correspond the covariances of conservation behavior and fitness (as measured by reproductive success) for both individuals and group averages (Fig. 5). Individual conservation-fitness covariance was nearly always negative in both treatments, because individuals who harvested at high rates could outcompete those who harvested at low rates, at least in the short run. By definition, there is no computable covariance by groups when only one group exists.

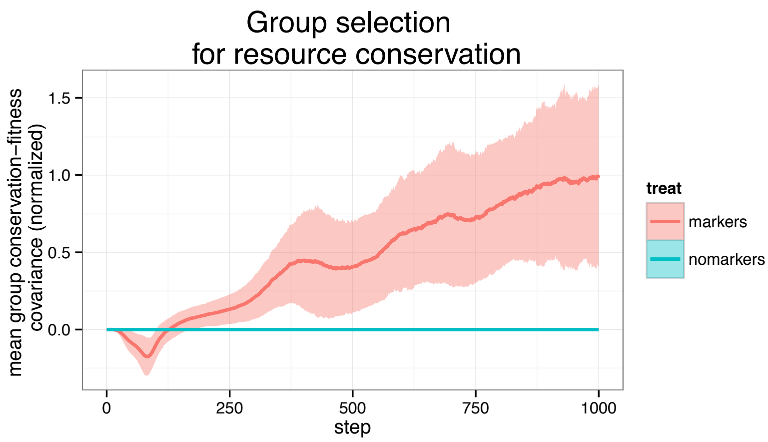

Fig. 5. Group selection of resource conservation (mean ±1 SD).

Selection for longer-lasting groups was responsible for the emergence of individual-level resource conservation and the supporting norms of production and boundary defense. To see this we plot the frequency with which populations persist to a given age as a function of group-level conservation-fitness covariance. Figure

6 demonstrates the relationship between group selection and population persistence, but also exposes key regions of instability or transition thresholds. These occur at $\texttt{cov}(w, c) = 0.0$, 0.4, 0.7, and 0.9. These transition regions occur when selection is strong (death rates are high) and populations are volatile (rapid population growth and decline). The first transition at $\texttt{cov}(w, c) = 0.0$ is the initial population crash and recovery, during which conservation behavior becomes strongly selected. The second and third transitions at $\texttt{cov}(w, c) = 0.4$ and 0.7 appear to occur when the populations pass through regions of strong selection, such as the evolution of a private property rights regime, and when cooperation norms degrade to “no one.”

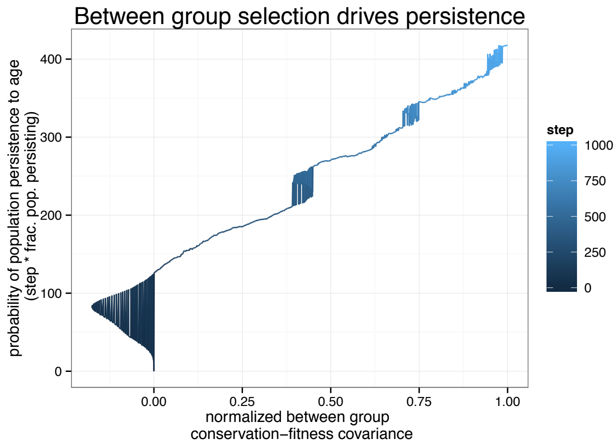

Fig. 6. Probability of population persistence to a given age as a function of the strength of group selection as measured by normalized between group conservation-fitness covariance, $\texttt{cov}(w, c)$.

3.3 Hypothesis 3: Social markers are required to potentiate group selection

Finally, we ran scenarios in which both social norms were allowed evolve with and without the presence of social markers. Price covariance measurements (H2) show us that group selection drives conservation and persistence. However, the Price breakdown cannot elucidate the specific effects of group markers, because with a single group, the between-group covariance of fitness and conservation behavior is zero by definition. To examine the effects of social markers on group-selected social norms, we therefore directly compared the persistence of populations in which social norms were allowed to evolve in the presence or absence of group markers (Fig. 7). Approximately 10% of populations survived without social marking. In these runs, populations were able to make a successful evolutionary transition to resource conservation. The availability of group markers, however, enabled much more rapid evolution of the same social norms of property and production, thereby allowing a much greater fraction of populations to survive.

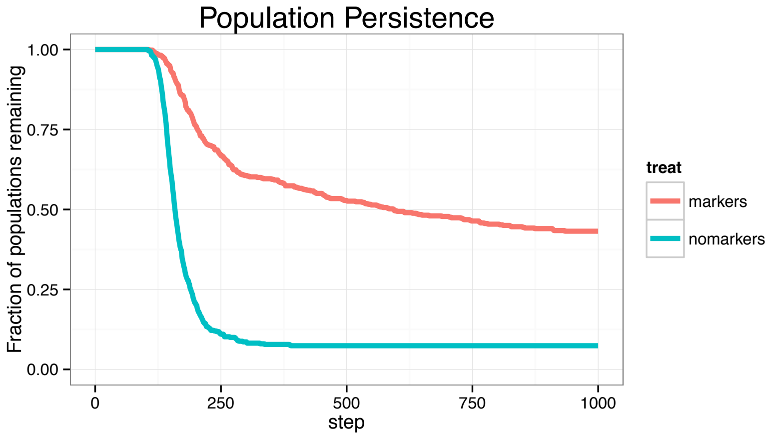

Fig. 7. Population persistence by the availability of social marking.

4 Conclusion

A model of endogenous group formation and evolution was built to examine the effects of group selection on individually costly resource conservation and related social norms. We have shown population persistence is dramatically enhanced by the endogenous emergence of two social norms, property defense and cooperative production, and that such an emergence occurs roughly 40% of the time. We then demonstrated that costly conservation behavior is favored by selection at the group level. Finally we have shown that the evolution of norms and institutions that promote cooperation – in the form of sustainable harvesting – is strongly facilitated by the presence of group markers that enable differential behavior toward in- and out-group individuals.

These findings have two major implications. First, our results show that evolutionary models that allow for group structure can elucidate mechanisms by which cooperative cultural institutions in general, and Ostrom's core design principles in particular, can emerge. The very first of Ostrom's principles for robust social-ecological systems management is that of usage boundaries on resource consumption. In our model, property regimes in the form of boundary defense emerge endogenously when social markers are allowed to evolve. When property regimes emerge, they dramatically enhance resource conservation and population sustainability. Our model therefore demonstrates that key social and institutional conditions for sustainability can be understood as the dynamic result of an ongoing process of cultural evolution at multiple organizational levels. It therefore stands to reason that the rest of Ostrom's principles may also be so modeled. This paper represents the first stage of a project aiming to do just that.

Secondly, our findings support an evolutionary theory of human sociality sometimes called the "tribal social instincts" model [35, 16]. The theory suggests that human social groups, and the psychological capacities that make social groups possible (ethnic marking, parochial altruism, etc.) evolved because they aided groups in solving collective action problems such as resource capture and management. Increasingly, research and theory suggest that our psychology was shaped by our extreme levels of sociality, particularly social organization into distinguishable groups [36, 37] . Our model shows that the ability to belong to a clearly delineated group – to distinguish individuals based on social cues and differentiate ones self with new social markers – is a feature that aids the evolution of institutions that solve the prisoner's dilemma of limited resource use. In fact our model supports the idea that group psychology – including cognitive capacities for cultural norms [17] such as property – enhance the evolution of collective action solutions.

References

1. Myers, R.A., Worm, B.: Rapid worldwide depletion of predatory fish communities. Nature **423** (2003) 280–283
2. Boyd, R., Richerson, P.J.: Culture and the evolutionary process. University of Chicago Press, Chicago, IL (1985)
3. Mesoudi, A.: Cultural evolution: How Darwinian theory can explain human culture and synthesize the social sciences. University of Chicago Press, Chicago, IL (2011)
4. North, D.C.: Institutions, institutional change and economic performance. Cambridge University Press, Cambridge, UK (1990)
5. Ostrom, E.: Governing the commons: The evolution of institutions for collective action. Cambridge University Press, Cambridge, UK (1990)
6. Wilson, D.S., Ostrom, E., Cox, M.E.: Generalizing the core design principles for the efficacy of groups. J Econ Behav Organ **in press** (2013)
7. Hardin, G.: The tragedy of the commons. Science **162** (1968) 1243–1248
8. Smith, J.M.: Evolution and the theory of games. Cambridge University Press, Cambridge, UK (1982)
9. Nowak, M.A., May, R.M.: Evolutionary games and spatial chaos. Nature **359** (1992) 826–829
10. Killingback, T., Bieri, J., Flatt, T.: Evolution in group-structured populations can resolve the tragedy of the commons. Proc R Soc Lond B **273** (2006) 1477–1481
11. Smaldino, P.E., Schank, J.C.: Movement patterns, social dynamics, and the evolution of cooperation. Theor Popul Biol **82** (2012) 48–58
12. Boyd, R., Richerson, P.J.: Punishment allows the evolution of cooperation (or anything else) in sizeable groups. Ethol Sociobiol **13** (1992) 171–195

13. Hauert, C., Traulsen, A., Brandt, H., Nowak, M.A., Sigmund, K.: Via freedom to coercion: The emergence of costly punishment. Science **316**(5833) (2007) 1905–1907
14. Nowak, M., Sigmund, K.: Evolution of indirect reciprocity. Nature **437** (2005) 1291–1298
15. Henrich, J.: Cultural group selection, coevolutionary processes and large-scale cooperation. J Econ Behav Organ **53** (2004) 3–35
16. Richerson, P.J., Boyd, R.: Not by genes alone: How culture transformed human evolution. University of Chicago Press, Chicago, IL (2005)
17. Chudek, M., Henrich, J.: Culture-gene coevolution, norm-psychology and the emergence of human prosociality. Trends Cogn Sci **15** (2011) 218–226
18. Wilson, D.S., Wilson, E.O.: Rethinking the theoretical foundation of sociobiology. Q Rev Biol **82** (2007) 327–348
19. Bell, A.V., Richerson, P.J., McElreath, R.: Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. Proc Natl Acad Sci USA **106** (2009) 17671–17674
20. Smaldino, P.E.: cultural evolution of emergent group-level traits. Behav Brain Sci **in press** (2013)
21. Bowles, S., Choi, J.K., .Hopfensitz, A.: The co-evolution of individual behaviors and social institutions. J Theor Biol **223** (2013) 135–147
22. van der Bergh, J.C.J.M., Gowdy, J.M.: A group selection perspective on economic behavior, institutions and organizations. J Econ Behav Organ **72** (20009) 1–20
23. Richerson, P.J., Henrich, J.: Tribal social instincts and the cultural evolution of institutions to solve collective action problems. Cliodynamics **3** (2012) 38–80
24. Winterhalder, B.: Diet choice, risk, and food sharing in a stochastic environment. J Anthropol Archaeol **in press** (1986) 451–463
25. Rousseau, J.J.: A discourse on inequality. Penguin Books, London (1755/1984)
26. Hammond, R.A., Axelrod, R.: The evolution of ethnocentrism. J Conflict Res **50** (2006) 926–936
27. McElreath, R., Boyd, R., Richerson, P.J.: Shared norms can lead to the evolution of ethnic markers. Curr Anthropol **44** (2003) 122–129
28. Bowles, S., Choi, J.K.: Coevolution of farming and private property during the early holocene. Proc Natl Acad Sci USA **in press** (2013)
29. Boyd, R., Gintis, H., Bowles, S., Richerson, P.J.: The evolution of altruistic punishment. Proc Natl Acad Sci USA **100** (2003) 3531–3535
30. García, J., van der Bergh, J.C.J.M.: Evolution of parochial altruism by multilevel selection. Evol Hum Behav **32** (2011) 277–287
31. Safarzynska, K.: The coevolution of culture and environment. J Theor Biol **322** (2013) 46–57
32. Smaldino, P.E., Schank, J.C., McElreath, R.: Increased costs of cooperation help cooperators in the long run. Am Nat **181** (2013) 451–463
33. Price, G.R.: Extensions of covariance selection mathematics. Annu Hum Genet Lond **35** (1972) 485–490
34. Okasha, S.: Evolution and the levels of selection. Oxford University Press, Oxford, UK (2006)
35. Richerson, P.J., Boyd, R.: The evolution of subjective commitment to groups: A tribal instincts hypothesis. In Neese, R., ed.: Evolution and The Capacity for Commitment, New York, Russell Sage Foundation (2001) 186–220
36. Caporael, L.R.: Coordinating bodies, minds and groups: Evolution and human social cognition. J Soc Evol Sys **19** (1996) 261–275

37. Brewer, M.B.: Taking the origins of human nature seriously: Towards a more imperialist social psychology. Pers Soc Psychol Rev **8** (2004) 107–113

Group Selection and Sustainability - ODD Protocol

1. Purpose

This model was constructed for the purpose of exploring the conditions under which sustainable management of a renewable and exhaustible natural resource is able to emerge within a simple society characterized by the potential for collaborative production for surplus and rudimentary group structure and mechanisms. Specifically, the model aims to simulate the emergence of key social and economic norms, Elinor Ostrom's Core Design Principles, commonly found in successful resource management institutions. The model allows social groups and economic regimes to emerge endogenously, and proliferate when successful, and fail due to resource collapse, to uncover those factors which increase the frequency of sustainable resource management.

2. Entities, state variables, and scales

Entities in the model include agents (individuals) and spatial units (patches of land). Although social groups exist, they do not act collectively, rather the agents' behavior is at times dependent upon group affiliations.

- *Agents.* The agents in the model have four primary traits which govern behavior; group identity, sharing trait, harvesting trait, and defense trait. In addition, agents are identified by the NetLogo program by a number, *who*, and their location is defined by the patch upon which they are standing. Agents are also characterized by their age, their lifetime reproductive success and the individual they choose to imitate. Three resource variables are associated with agents; a running total of the resources accumulated within a time step, as well as two accumulated resource totals, raw and processed. Table 1 lists the state variables and provides a brief description of each.

Table 1. Agent variables

Variable	*Description*	*Bounds*
Group identity, g_i	Marker that identifies agent as a member of a given social group	[1,9]
Sharing trait, s_i	Ternary trait determining agent's behavior in a pairwise sharing encounter: share with no one (N), group members (G), or all individuals (A)	A,N,G
Harvesting trait, h_i	Binary trait (*high harvesters* and *low harvesters*) determining maximum raw resources an agent attempts to harvest from its available land	H,L
Defense trait, d_i	Ternary trait determining which neighboring agents an agent will defend its property (and, specifically, its harvestable resources) against: defend against no one (N), out-group members (O), or all individuals (A)	A,N,O

Agent identity, who_i	Identifier associated with the agent (internal to NetLogo)	positive integer
Agent location, $patch\_at_i$	Location associated with the agent (internal to NetLogo) as defined by the patch the agent is currently standing on	(x,y) integer pair
Agent age, age_i	Number of simulation time steps the agent has survived	positive integer
Harvested amount, H_i	Running total of the amount of resource that has been harvested at a given time step	positive integer
Raw resources, r_i	Accumulated resources harvested from the land	positive integer
Processed resources, p_i	Accumulated resources produced by contributing raw resources to a cooperative process (modeled as a two-person public goods game)	positive integer
Agent to copy, $copy_i$	Neighboring individual that the agent has chosen to imitate within the time step	positive integer
Reproductive success, rs_i	Running total of the number of offspring of an agent within its lifespan	positive integer

– *Patches.* These entities have only two attributes, patch resources and patch regrowth rate. These are listed in Table 2 below along with their descriptions.

Table 2. Patches and their state variables

Variable	*Description*	*Bounds*
Patch resources, $crop_j$	Amount of resource currently growing upon the patch	positive integer
Patch regrowth, pr_j	Local patch maximum intrinsic rate of growth	positive integer

– *Environment.* The model has no externally varying environmental conditions.

– *Collectives*. Although agents have a group identity, this group does not act as a collective. No joint decisions are made, and although social identity is shared, all actions are individual.

– *Spatial and Temporal Scales.* Each grid cell or patch represents a one hectare plot of land. The model is a 32 x 32 torus. Each time step represents approximately one year.

3. Process overview and scheduling

The following provides a general version of the process overview and scheduling followed by pseudo-code detail. For more detail regarding each of the procedures, please see the *submodels* section.

1. *Set up model world:* Agents, patches and globals are set up.
2. *Patch defense*: Agents pay a cost to defend their focal patch.
3. *Harvesting*: Agents harvest raw resources from their focal and neighboring patches.
4. *Sharing*: Agents have the opportunity to collaborate with a neighbor to produce added-value processed resources.
5. *Pay cost of living*: Agents lose resources in order to stay alive.
6. *Death*: Agents die if their resources drop below zero or due to a random events, the probability of which increases with age.
7. *Reproduction*: Agents with sufficient resources attempt to reproduce.
8. *Migration*: Agents move to an empty neighboring patch.
9. *Imitation*: Agents may imitate the traits of others with certain biases.
10. *Patch growth*: Patch resources increase if they are below maximum.
11. *Aging*: Agents grow older.
12. *Cap resources*: Agents cannot hold resources above a threshold between rounds.

– *Process overview and scheduling, pseudo-code.*

1. Setup the model world

- clear all entities and variables
- set up the treatments for testing hypotheses (set global parameters)
- set up the remaining global parameters that are not set within the treatments set the amount of resources on the patches to some random number between 0 and the carrying capacity
- set all patches to initially be unowned/undefended
- color the patches according to the amount of resource on the patch (for visualization purposes only)
- create an agent for each combination of defense trait and share trait
- have each agent create n-1 more agents where n is the desired total number of agents per group
- move these new agents to the closest patch to the group-creating agent that is empty
- apply treatment conditions to agent traits
- set all of the agents to be low harvesters
- set the agents' initial processed resources to a random amount between 0 and the cost of reproduction
- if the agent is a low harvester
    - set the agent's shape to be a circle
    - otherwise, set the agent's shape to be a square
- reset the time step counter

2. Go/actions - repeated until simulation is aborted

- if there are no agents left in the world
    - stop the simulation

- agents choose to defend their property based upon their defense trait and pay cost of defense
- agents harvest their preferred amount based upon their harvest trait
- if there is another individual in the agent's Moore neighborhood
    - play a public goods game with a random neighbor
- agents' resources, starting with processed resources, are reduced by the cost of living
- if agent's resources are less than 0 or stochastic age-based expression is true
    - agent dies
- if agent's total resources exceed the cost of reproduction and there is an empty patch in the Moore neighborhood
    - agent creates an offspring
- if a random number between 0 and 99 is less than the probability of migration and there is an empty patch in the Moore neighborhood
    - agent migrates to the patch in the Moore neighborhood with the most resource
- if a random number between 0 and 99 is less than the probability of imitation and there is another individual available
    - imitate an individual according to imitate trait
- if the resources on the patch are greater than zero
- regrow resource according to logistic growth function subject to a given carrying capacity
    - otherwise, allow a very small chance of regrowth on an empty patch
- increase agents' age by 1
- reduce agents' resources to comply with resource caps
- recolor patches based upon the amount of resource remaining
- if the agent is a low harvester
    - set the agent's shape to be a circle
    - otherwise, set the agent's shape to be a square
- advance to next time step

4. Design concepts

Basic principles. This model places a spatially-explicit ABM of the evolution of cooperation within the context of a social-ecological system. It draws upon theory regarding cooperation, commons dilemmas, group selection, and Elinor Ostrom's Core Design Principles.

Emergence. Spatially coherent social groups of like agents emerge endogenously. Groups that persist sometimes consist of "groupish" and cooperative agents who harvest at low levels.

Adaptation. Agents adapt at both genetic and cultural levels. Agents have a simple 4 locus 'behaviorome' [share, harvest, defense, marker]. In the analog of genetic adaptation, mutation of all these traits occurs with some given probability at reproduction. Death and differential reproduction then create selection and adaptation on these loci over time. In the cultural analog, agents adapt every round through a process of payoff-biased imitation.

Objective. Agents do not have internalized goal states, or encoded objectives.

Learning. There is no learning in this model.

Prediction. There is no prediction in this model.

Sensing. Agents can sense the group marker of the other agents in its neighborhood or within a certain radius of their patch. Agents are also able to sense the amount of resource on their current patch as well as the amount of resource on neighboring patches. At times, agents are able to observe the wealth of neighbors.

Interaction. Agents interact with patches by harvesting resources. Agents interact with other agents through defense of their patch, engagement in cooperative production and imitation.

Stochasticity. There is stochasticity in the initialization of patch resources, in the placement of the initial agents, in the assignment of traits to agents, in migration, imitation, and mutation of offspring traits, and in all probabilistic procedures.

Collectives. Each agent-agent interaction for the purposes of cooperative production can be seen as a brief collective in that the payoff received from the public goods game is dependent upon the participation of both actors and is split between the two parties. The collective then splits at the end of the round. Alternatively socially-marked groups can be seen as collectives, with common social marking, (and, depending on the group) common sharing, and harvesting traits. Collectives do not however perform collective behaviors *per se*. Each Moore neighborhood of patches is a resource collective, a commons. Agents are able to harvest from any of the patches in their commons subject to the defense traits of any owners of those patches.

Observation. We collected population size, frequency of low harvest norm, frequencies of each production norm, frequencies of each property norm, number of groups, mean individual conservation-fitness covariance, and group conservation-fitness covariance by timestep.

5. Initialization

The model world is initialized with nine groups of 12 agents each. Groups are located randomly on the 32x32 patch grid. All individuals begin the simulations with the conservative harvesting trait.

The following global parameters, listed below in Table 3, are set at model initialization. The values at which they are set within the model are detailed in section eight, *Simulation experiments/model analysis*.

Table 3. Global parameters

Parameter	*Symbol*	*Description*	*Value*
Max growth rate	r	The maximum rate of (logistic) growth for the patch resources	0.5
Carrying capacity	K	The maximum resources a patch can contain	200
Lattice size	L	The width and length of the square lattice	32 x 32
numGroups	G_{init}	The initial number of groups.	9
Harvest gap	G_H	The proportion of MSY by which high harvesting amount exceeds low harvest amount	1
Sharing proportion	γ	The maximum proportion of an agent's raw resources it contributes to a pairwise public goods game.	0.5
Marginal defense cost	C_D	The cost of defending a patch for each neighbor an agent defends against	1
Imitation proba-	λ	The probability that an agent imitates a trait of another agent.	5%

bility		Independent for each trait (sharing, harvesting, defense).	(0.05)
Imitation radius		The radius (neighborhood size) within which models for imitation can be found.	2
Public goods production factor	Θ	The amount the raw resources contributed to the public goods game are multiplied by to yield the payoffs in processed resources. Constraint: $1 < \Theta < 2$.	1.5
Mutation rate	μ	The independent probability that each offspring trait is not copied from the parent but is randomly selected.	0.3% (0.003)
Migration rate	m	The probability that an agent attempts to move to a neighboring patch. Set to zero in the current model.	0
Mutate markers		Boolean. If true, agents may also mutate their group markers.	[0,1]
Raw resource cap		The maximum amount of raw resources an agent can accumulate	1,000,000
Processed resource cap		The maximum amount of processed resources an agent can accumulate	1,000,000
Production		Boolean. If true, cooperative production may take on values other than the base institution (no production).	[0,1]
Property		Boolean. If true, property defense may take on values other than the base institution (no defense).	[0,1]
Markers		Boolean. If true, agents may take on markers values that differentiate them from other agents.	[0,1]

Calculated Parameters		*Description*	*Calculation*
Maximum sustainable yield	*MSY*	Maximum sustainable yield.	floor ((K * r) / 4)
High harvesting rate	H_H	The maximum amount of raw resources a high harvester attempts to extract from its commons.	floor $((1 - (G_H / 2)) * MSY)$
Low harvesting rate	H_L	The maximum amount of raw resources a low harvester attempts to extract from its commons.	floor $((1 + (G_H / 2)) * MSY)$
Cost of living	C_L	The amount of resources lost by an agent each time step. Agents can use either raw or processed resources to pay this cost.	floor (0.2 * MSY)
Cost of reproduction	C_R	The amount of resources an agent loses when it reproduces.	floor $(2 * (H_H - C_L))$

6. Input data

This model does not use input data to represent time-varying processes or spatial patterns.

7. Submodels

Patch defense – Agents pay a cost to defend their focal patch.

- If $d_i = N$, the agent does not pay a cost and the patch is undefended.
- If $d_i = O$, the agent pays C_D for each out-group agent in its neighborhood. Those agents are then prevented from harvesting on the agent's focal patch this time step.
- If $d_i = A$, the agent pays C_D for each agent in its neighborhood (including in-group neighbors). Those agents are then prevented from harvesting on the agent's focal patch this time step.
- If the agent does not have sufficient resources to pay its costs, it does not pay them and its patch can be harvested by neighbors.

- Costs are deducted first from processed resources, and are only deducted from raw resources once processed resources are depleted.

Harvesting – Agents harvest raw resources from their focal and neighboring patches.

- Agents are able to harvest (defined as decreasing a patch's *patchResource* by some amount and adding the same amount to the agent's agentWealth) from their *commons*, defined as their focal patch plus the patches in their Moore neighborhood, excluding any patches that are defended against them.
- This procedure starts by asking the agent to add up all the resources in their commons. If this amount is less than the agent's harvest amount as determined by their *harvesting trait*, they harvest the entire commons by adding the resources to their raw resources and setting *patchResource* of all of the patches in their commons to 0.
- If there are more resources in the commons than they need, they start by harvesting from the patch with the most resources and continue harvesting from decreasingly plentiful patches until they have harvested the desired amount of resource.
- This procedure relies upon NetLogo's internal scheduling capabilities to ensure that the same agents do not always harvest first, thereby avoiding any resource harvesting bias.

Sharing – Agents have the opportunity to collaborate with a neighbor to produce added-value processed resources.

- Each agent executes the following process in random order.
- An agent chooses a random agent in their neighborhood (if any), and the two play a two-person public goods game.
- An agent's contribution x_i is equal to zero if the agent shares with no one (N) or if the agent shares with in-group members (G) and its partner is not a member of its group. Otherwise, the agent contributes a proportion of its raw resources. An agent's contribution is calculated using the following algorithm:
    - **if** $p_i + r_i < C_L$
        - $x_i = 0$
    - **else if** $p_i + (1 - \gamma)r_i \geq C_L$
        - $x_i = \gamma r_i$
    - **else**
        - $x_i = p_i + r_i - C_L$
- If an agent has insufficient resources, it only contributes the fraction of its raw resources such that it retains enough total resources to pay the cost of living.
- The total contribution of both agents is multiplied by a factor Θ and divided evenly between both agents.

Pay cost of living – Agents lose resources in order to stay alive.

- Each agent loses C_L resources. Resources are first taken from processed resources, and then raw resources if the agent does not have enough processed resources.

Death – Agents die if their resources drop below zero or due to a random events, the probability of which increases with age.

- Agents are removed from the simulation if either
    - Their total resources are less than zero.
    - They randomly die as per an age-linked function:
      $\Pr(die) = \frac{1}{1+\exp\,[10-\frac{age}{8}]}$

Reproduction – Agents with sufficient resources attempt to reproduce.

- If an agent has total resources greater than or equal to the cost of reproduction (i.e., $p_i + r_i \geq C_R$) and if there is an empty space somewhere in the agent's Moore neighborhood, the agent loses the cost of reproduction (with deductions from processed resources first) and produces an offspring on the chosen patch.
- The offspring's initial resources are set to zero.
- The offspring's sharing, harvesting, and defense traits are each identical to the parent's with probability $1 - \mu$, and are randomly selected otherwise. Mutation for each trait is independent.

Migration – Agents move to an empty neighboring patch.

- With probability m, an agent attempts to move.
- The agent selects a random empty patch in its Moore neighborhood (if any), and moves there.
- In the current model, $m = 0$.

Imitation – Agents may imitate the traits of others with certain biases.

- An agent decides to imitate the trait(s) of another agent with probability λ.
- If so, the agent looks among the other agents in its *imitation radius* for a model to imitate.
- If the *imitation treatment* is group (G), the model must also be chosen from among only those agents who have the focal agent's group identity. Otherwise, if the imitation treatment is all (A), the model can be chosen from among all agents in the imitation radius.
- A model is chosen using one of the following *imitation biases*:
    - Payoff (P): The model is the agent with the most total (raw + processed) resources.
    - Green (G): The model is chosen randomly from among those agents with a *low* harvesting trait.
    - Random (R): The model is chosen randomly.
    - Conformity (C): The most common behavior among the other agents in its *imitation radius* is selected.
        - These imitation biases may be merged.
- Depending on the global parameters, any or all of the following traits – harvesting, sharing, and group identity – may be imitated.

Patch growth – Patch resources increase if they are below maximum.

- If the patch resources are nonzero, they grow according to a density-dependent logistic function:
    - $\pi_{t+1} = \pi_t \left(1 - r\left(1 - \frac{\pi_t}{K}\right)\right)$
    - where r is the maximum intrinsic rate of growth and K is the carrying capacity.
- Otherwise, if the patch resources are zero (i.e., they have been completely depleted, they are set to 3 (a low number) with probability of 0.001.

Aging – Agents grow older.

- Agents increase their age by 1.
- This only affects their probability of dying randomly.

Cap resources

– The raw and/or processed resources are capped at their maxima each time step.

8. Simulation experiments/model analysis

We tested our hypotheses with a set of five experimental treatments. These five treatments vary three parameters, each of which represents the possibility of endogenous cultural evolution in a certain type of individual trait. The three traits are: production norms, property norms, and social markers. These parameters are combined in treatments as follows:

Table 4. Treatment conditions

Treatment	*Name*	**Production**	**Property**	**Markers**	*Hypothesis*
1	no norms			Y	1
2	no prop	Y		Y	1
3	no prod		Y	Y	1
4	both norms	Y	Y	Y	1,2,3
5	no markers	Y	Y		3

These treatments were applied to test the three hypotheses as follows

1. *Hypothesis 1*: When conservation is individually costly, population persistence and sustainable resource management requires social norms or institutional features, including shared production and boundary defense.
   - Treatments 1,2,3,4
2. *Hypothesis 2:* When resource conservation, sharing, and boundary defense are individually costly, they should not be stable without selection acting at the level of the social group.
   - Treatment 4, within and between group covariance
3. *Hypothesis 3:* For group-level selection to occur, social group markers must be present to promote preferential behavior toward group members.
   - Treatment 4,5

All individuals begin the simulations with the conservative harvesting trait, and social norms are initialized so that individuals are homogenous within groups. When markers are present, groups begin with different markers, which may then evolve via imitation and mutation. When social marking is disabled, all individuals have a single marker trait that does not vary. Likewise, when either production or property norms are active, one entire group is initialized with each unique combination of norms, so that initial populations are always identical. For instance, if only one norm type is active, for instance, then three groups will be initialized with each of the three property norm states. When both norms are active, then one group will be initialized with each of the nine combinations of states of the two norms.

Global variable values used within the treatments are as detailed in Table 3.