Title

Ratchet or hatchet? Modeling the cultural evolution of simplified construals

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors
Roberts-Gaal, Xavier
Callaway, Frederick
Cushman, Fiery

Publication Date
2026

Ratchet or hatchet? Modeling the cultural evolution of simplified construals

Xavier Roberts-Gaal[1], Frederick Callaway[1,2] & Fiery Cushman[1]
[1]Department of Psychology, Harvard University
[2]Department Psychology, New York University

Abstract

Cumulative culture is often described as a ratchet that builds ever more complex structures, knowledge, and technology. However, this complexity comes at a cost. Though cultural innovations may offer new opportunities or insights, they also present new ways to become confused or cognitively burdened. Cultural evolution may thus act both like a ratchet (accumulating useful innovations) and like a hatchet (stripping away representational deadwood). Building on a recent theory of value-guided construal, we formalize this utility-complexity trade-off using a model inspired by navigating with a map. We embed this model in an evolutionary dynamic, where the maps of successful navigators are selectively copied. As expected, more complex (simpler) maps are favored when they are more useful and when agents have a higher (lower) representational capacity. Yet surprisingly, the evolutionarily stable map is often more or less complex than the optimal map, driven by the availability of social information. More broadly, our modeling results show how cognitive costs, social learning, and cultural selection can jointly shape our minds and the material artifacts that aid them.

Keywords: task construal, cultural evolution, resource-rationality, planning, maps

Introduction

Cumulative culture is a hallmark of our species' success (Henrich, 2017; Tennie et al., 2009). As cultures accumulate adaptive modifications, their artifacts, institutions, and practices 'ratchet,' becoming increasingly powerful and sophisticated and attaining levels of complexity that no single person can fully understand or reinvent from scratch (Chater & Christiansen, 2025; Enquist et al., 2008). Like a ratchet that can move in only one direction, it is proposed, the complexity of cultural forms compounds with each generational tweak. Cultural ratcheting has been observed in many domains where human group-specific behavior emerges, from material artifacts and technology (e.g., composite bows), to institutions (e.g., administrative bureaucracy), to procedures, traditions, and beliefs (e.g., traditional processing of cassava).

Yet, we face cognitive constraints that limit our use of such complex and sophisticated cultural machinery, as decades of value-guided decision-making research underscores (Lieder et al., 2026). Accordingly, we often rely upon approximate rather than precise representations of problems. One example is *task construals*, which frame planning problems in simpler terms (Ho et al., 2022). Like all mental models, approximations may be culturally transmitted via social learning (Wu et al., 2022) or embodied in artifacts, such as maps.

Intuitively, as cultures gain the capability to accomplish more complex tasks in the world, they should require increasingly complex representations. And as a culture's territory expands, its maps should grow to represent more information. But, as Borges' imperial cartographers realized, a map as detailed as the territory it represents is useless. The best map, then, leaves out some details—shaping how we think about the problem of navigation by focusing attention on certain routes to the exclusion of others. A simpler model often can be more effective than a complex model, because it is easier to use, even if the complex model is ultimately more accurate. (Kepler's celestial mechanics and Newton's laws of motion, for example, are perfectly adequate for predicting eclipses, and unlike general relativity, do not require solving partial differential equations.) A natural way of expressing this advantage of simplicity is in terms of the *computational cost* of problem-solving with a model. Like a map at the right level of detail, the optimal model of a problem distills its problem's essential features, striking a balance between the cost of using that representation to compute a solution and the quality of the resulting solution.

Although the accumulating complexity of culturally transmitted models has received much attention, the role of simplicity has been comparatively under-studied. In particular, the circumstances in which cultural selection favors simpler models over complex ones—acting as a 'cultural hatchet'—have not been fully characterized. Here, we develop and explore a computational model that encapsulates the core idea that sometimes cultural selection should exert a simplifying force on socially transmitted ideas, representations, and artifacts.

This raises a crucial question: When will cultural evolution accumulate complexity (the 'cultural ratchet'), and when will it instead exert a force toward simplicity (a 'cultural hatchet')? To address this, we consider a minimal model of cultural evolution in the context of construal-guided navigation. The core insight contained in this model is that two kinds of resource limitation—the computational cost of planning and the sample scarcity of social learning—govern when cultural evolution accumulates complexity, and when instead it simplifies.

Background

Value-guided task construals

Our model is built upon the framework of *value-guided task construals*, adopting the approach taken by Ho et al. (2022). We consider an agent solving a Markov Decision Process

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

(MDP) $\mathcal{M}$, consisting of a state space $\mathcal{S}$, an action space $\mathcal{A}$, a transition function $P : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$, and a utility function $U : \mathcal{S} \to \mathbb{R}$. To solve the MDP, an agent optimizes a policy $\pi : \mathcal{S} \times \mathcal{A} \to [0, 1]$, whose value V_π is determined by the cumulative utility U of acting in accordance with it from the starting state s_0:

$$V_\pi = U(s_0) + \sum_a \pi(a \mid s_0) \sum_{s'} P(s' \mid s_0, a) V_\pi(s').$$

Of course, the agent could attempt to find the highest-value policy by planning over their knowledge of the true transition function, at significant cognitive cost (e.g., a lengthy decision time bearing an opportunity cost). But, another option for the agent is to select a simplified representation of the problem for which it can efficiently compute the optimal policy (which, ideally, will achieve an outcome close to the optimal policy for the full problem). That is, the agent could construe a *modified* MDP $\mathcal{M}' \subseteq \mathcal{M}$, potentially omitting some of the ground-truth MDP's state, action space, or transition function. Because the construal simplifies the actual transition dynamics, the agent can compute the optimal policy according to the construal, π_m, with lesser cognitive cost. This resulting policy may differ from the one the agent would have computed had they taken the full transition function P into account (potentially to the agent's detriment). Its utility is given by its interactions with the ground-truth dynamics:

$$U(\pi_m) = U(s_0) + \sum_a \pi_m(a \mid s_0) \sum_{s'} P(s' \mid s_0, a) V_{\pi_m}(s').$$

This formalization allows us to consider the *value of representing* a construal (VOR). As in models of resource-rationality (Callaway et al., 2022) and the expected value of control (Shenhav et al., 2013), we penalize the utility associated with a construal by the cognitive cost C of using that construal to plan. A straightforward way to do this, with minimal assumptions about the form of the planning algorithm used by the agent, is to penalize the *complexity* of the construal $C(\mathcal{M}') = |\mathcal{M}'|$ (i.e., its size, see Ho et al., 2022). The value of having a construal is then its utility minus its cognitive cost, scaled by a parameter ψ:

$$\text{VOR}(m) = U(\pi_m) - \psi C(m).$$

To summarize, we rely upon the notion of a task construal introduced in Ho et al. (2022). This involves relaxing the assumption that an agent must plan over the full environment in which they act; instead, the agent can omit some components from its model of the environment to more efficiently compute a policy. The optimal construal is one that appropriately trades off the cost (complexity) of the construal, and the real-world utility of the policy the agent ultimately computes using that construal.

Models of cultural evolution

Turning to the cultural level, various approaches have been proposed to model cultural evolution, the process by which agents acquire knowledge or skills from forebears and transmit them to the next generation.

We follow a standard formalization, defining a culture as a (potentially infinite) population of agents, each belonging to one cultural variant. (For instance, the variants could be different construals of an MDP.) Each cultural variant is associated with a fitness distribution. Cultural variants with higher fitness in one generation are more likely to be transmitted to agents in subsequent generations. To model this phenomenon, we rely upon the insight that cultural transmission occurs through *social learning* mechanisms, which are often payoff-biased (Barrett et al., 2017; Kendal et al., 2009; McElreath et al., 2008). Moreover, although social learning allows individuals to cheaply acquire knowledge from others in their culture, it is not an unlimited channel. Individuals can only pay attention to a certain number of teachers, and are constrained by the availability of social interaction with high-fitness ancestors. Specifically, we will adopt and generalize a common way of modeling the payoff- and frequency-sensitive nature of social learning, best-of-k learning (Baldini, 2013), in which individuals pay attention to k potential teachers and copy the variant of the highest-fitness teacher.

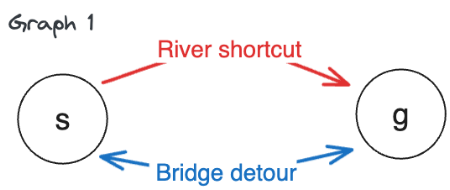
Computational model

With these preliminaries in place, we now introduce a minimal model of the cultural evolution of task construals. Specifically, we look at the context of navigating using a map. We use this computational model to show when cultural evolution ratchets and when it instead hatchets complexity.

Of course, if the population is entirely homogeneous and fitness is deterministic, then culture (like any other optimizer) is likely to converge on the optimal construal: one that strikes the fitness-maximizing balance between utility and complexity, which is the same for all agents. Thus, an important element of our modeling strategy is to consider *individual variability* in how different maps confer benefits and, therefore, in the optimal construal. We will consider two such ways in which maps may confer varying benefits, beginning with the idea of a an environment that contains a route that is potentially advantageous to some individuals, but not to others. We explore the conditions governing when this route tends to be represented in the culture's maps.

Setting

We consider the task of navigating on a directed graph $\mathcal{G} = (\mathcal{S}, \mathcal{E})$, illustrated in Figure 1. Specifically, the agent must start at an initial state $s \in \mathcal{S}$, visit a goal state $g \in \mathcal{S}$, and return to the initial state s. To make this concrete, imagine that agents have to navigate from their village to the market and then return home. Two edges $e \in \mathcal{E}$ connect s to g: a shortcut (e.g., swimming downstream to reach the market quickly), and a detour (e.g., following a circuitous route across a bridge). We denote these edges as $s \overset{river}{\to} g$ and $s \overset{bridge}{\leftrightarrow} g$, respectively. Crucially, while the bridge is bidirectional, the river can only be taken in one direction (downstream).



Utilities	Swimmer	Non-swimmer
River shortcut	-1	-∞
Bridge detour	-2	-2

Figure 1: Graph with a shortcut route. Agents start at s and must journey to g. They can always take a longer route via the bridge, but in one direction, they can take a shortcut via the river. Only agents who can swim benefit from taking the river shortcut. Utilities for each action, depending on the agent's type τ, are given in the table.

To model the effects of population diversity, we additionally endow agents with a binary type τ that determines the utility associated with traversing each edge, i.e. $U(e, \tau)$. Concretely, we take $\tau \in \{\text{swimmer}, \neg\text{swimmer}\}$ and assume that attempting to float downstream without knowing how to swim leads to a swift and certain death (see Figure 1).

These relatively minimal assumptions—the presence of a sometimes-useful shortcut that provides differential access to its benefits—allow us to derive a normative account of how agents should represent the graph $\mathcal{G}$.

Individual planning with a construal

Each agent uses a *map* $m \subseteq \mathcal{E}$ to help them plan routes to their goal (i.e., a subgraph that serves as a task construal). In principle, there are four possible maps corresponding to the powerset of edges. However, because all agents must take the bridge on the return trip, maps without the bridge would be strongly selected against; we thus ignore them for simplicity.[1] This leaves two possible maps: one that includes only the bridge detour ($m_1 = \{s \overset{bridge}{\leftrightarrow} g\}$) and a second that includes both the bridge detour and the river shortcut ($m_2 = \{s \overset{bridge}{\leftrightarrow} g, s \overset{river}{\rightarrow} g\}$).

Clearly, the full map provides higher average utility because it allows swimmers to take advantage of the shortcut. However, computing the shortest journey over the full map requires considering more potential routes, thus incurring an additional cognitive cost compared to using the partial map containing only the detour. Following Ho et al. (2022), we assume that the cost of a construal is proportional to its size; here, we set the cost to the number of edges included in the map, $C(m) = |m|$, and scale cost by a parameter ψ. Moreover, although only swimmers can take advantage of the shortcut, all agents pay the cost of considering the route. To account for this, we extend the value of representation (VOR) to account for the agent's type as well as their construal (c.f. the task-modulated VOR used in Ho et al., 2023):

$$\text{VOR}(m, \tau) = U(\pi_{m,\tau}) - \psi C(m).$$

Thus, the value of representing of each map is a function of its type-dependent utility (determined by whether one can take advantage of the shortcut when it is available) and its cognitive cost (determined by the size of the agent's map).

Cultural evolution of construals

Agents inherit their swimming skill with some probability Pr(swimmer). However, an agent acquires their map from an agent in the previous generation through social learning, observing the maps and fitness of agents from the previous generation. We assume that observed fitness is equal to the VOR for the demonstrator's construal and type, corrupted by Gaussian noise with variance σ^2. This captures variance in fitness due to unrelated factors and/or imperfections in the selection process itself. Under these assumptions, the fitness distribution X_i for a given construal i is a mixture of normals, with one component for each type τ:

$$X_i \mid \tau \sim \text{Normal}(\text{VOR}(m_i, \tau), \sigma^2).$$

For analytic tractability, we adopt a generalization of best-of-k learning (Baldini, 2013) in which the number of teachers each agent has access to, k, is drawn from a Poisson distribution: $k \sim \text{Poisson}(\lambda)$. This corresponds to assuming that every agent observes a new teacher at a constant rate λ for a fixed period of time before they commit to using one of their maps. Agents prefer to copy successful agents, so they adopt the variant of the best-performing agent they observe. The probability that a given agent copies map i (which we denote q_i) thus depends on three things: the proportion of other agents using the map (p_i), the fitness distribution for the map (X_i), and the fitness distribution for competing maps (p_j and X_j for $j \neq i$). Concretely, the probability of copying a map is defined:

$$q_i = \int_{-\infty}^{\infty} \lambda p_i f_i(x) \exp\Big(-\sum_j \lambda p_j \left[1 - F_j(x)\right]\Big)\, dx, \qquad (1)$$

where $f_i(x)$ and $F_i(x)$ are the density and cumulative density of the fitness of construal i (distributed according to X_i above). Intuitively, we integrate over possible maximal fitness values x (i.e., the maximum fitness value, drawn from X_i, observed among a set of sampled teachers) and assign density to a map i according to the probability that (1) a teacher using it is sampled, λp_i, (2) that teacher has the maximal fitness value, $f_i(x)$, and (3) no other teacher in the set has a map with a higher value (the exponentiated term).

Note that the interaction rate λ controls the strength of selection; when λ is low, agents may fail to observe all the cultural variants, and may therefore copy a construal that has less fitness than the maximum-fitness construal in the population.

[1]Practically, including these maps would merely reduce the rate of social learning by a small amount determined by the mutation rate.

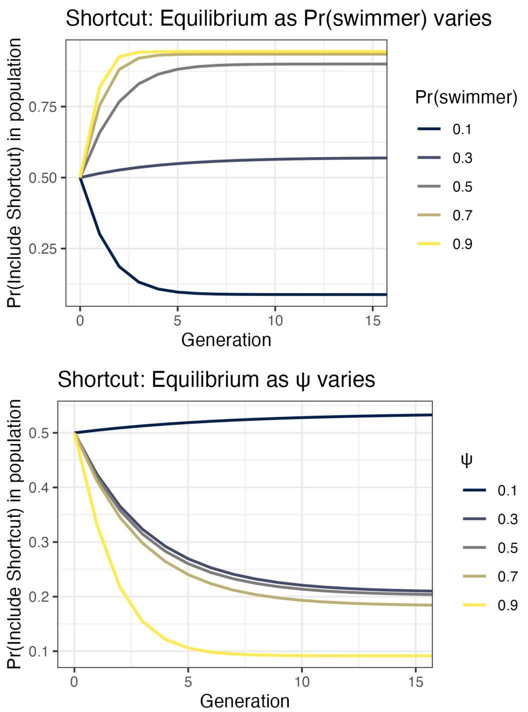


Figure 2: **Top:** Proportion of maps that include the shortcut over successive generations, split by the probability of benefiting from the shortcut, Pr(swimmer), fixing $\psi = 0.5$. This chart fixes Pr(swimmer = 0.2). **Bottom:** The same, split by computational cost ψ, which penalizes including the shortcut.

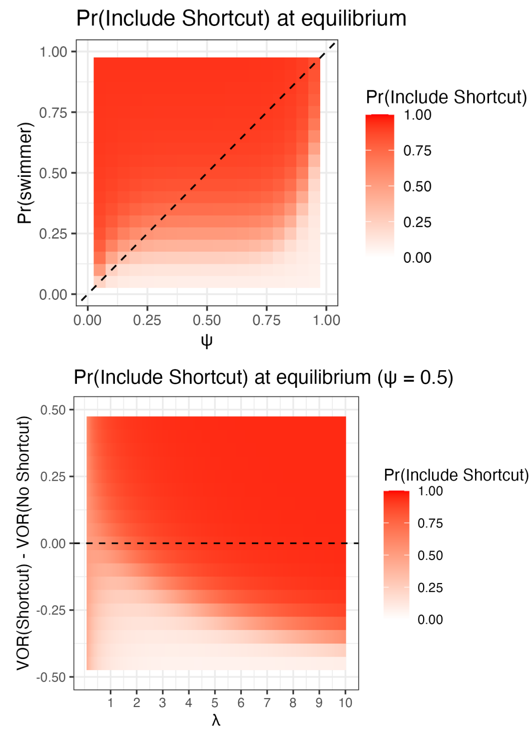


Figure 3: **Top:** The equilibrium probability the culturally evolved map includes the shortcut, as a function of these two parameters (fixing $\lambda = 7$). **Bottom:** Payoff-biased social learning shifts the cultural evolutionary equilibrium toward including the shortcut even when doing so is detrimental to mean population fitness (i.e., the VOR of the map including the shortcut is less than the VOR of the map without). This bias gets stronger as social learning becomes more plentiful (i.e., the interaction rate λ increases).

As $\lambda \to \infty$, agents become virtually guaranteed to observe all variants and copy the overall best in the population.

As is standard in (cultural) evolutionary models, we additionally assume a small probability η of random mutation between cultural variants (e.g., due to copying errors). The complete evolutionary dynamics for generation t are thus given by

$$p_i(t+1) = (1-\eta)q_i(t) + \frac{\eta}{N_m - 1}(1 - q_i(t)),$$

where N_m is the number of construals (in our case, $N_m = 2$).

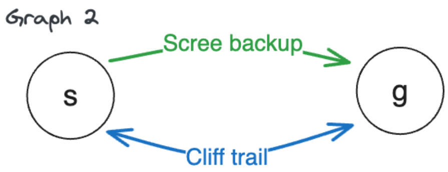
In this setting, we ask: When does cultural evolution favor the complexity by representing the river shortcut, and when does it instead favor simplicity by omitting it? To answer this, we explore how three parameters determine what cultural evolution favors: (1) the probability of belonging to each type (e.g., Pr(swimmer)) and (2) the scalar importance of cognitive cost ψ, which together control the value of construals, and (3) the interaction rate λ, which affects the availability of social information. (For simplicity, we fix the two nuisance parameters, fitness noise (σ) and mutation rate (η), to 0.1 and 0.05, respectively, but our results are not sensitive to these precise values.) We now turn to analyzing this model.

Results: Computational model

Inspecting the model, several results are immediately apparent (Figure 2). First, and as one might expect, when the expected utility of having the river shortcut in one's map is high (because more agents can take advantage of it), maps with the shortcut become more frequent over successive generations. Second, when the representational cost is weighed more heavily (e.g., because the agents are under time pressure or are not especially bright), maps with the shortcut become less frequent over successive generations.

Of course, these two parameters trade off, as shown in the top of Figure 3. However, the region in which the shortcut is kept extends well beyond the population-optimal indifference point (when the probability of saving 1 util by taking the shortcut equals the representational cost). That is, even when cultural evolution "should" be acting as a hatchet, it continues to favor the more complex map.

This mismatch between the globally optimal construal and the one selected by cultural evolution is more clearly illustrated in the bottom panel of Figure 3, where we have transformed the y axis to show the relative value of the complex construal. Even more surprisingly, we see that this bias becomes increasingly strong as agents have access to more social information. To see why, consider that representing the shortcut provides a



Utilities	Acrophobe	Non-A'phobe
Scree backup	-2	-2
Cliff trail	-3	-2

Figure 4: Graph with a backup route. As before, agents start at s and must journey to g. But they can either take a scary cliff edge trail (less useful to those agents with fear of heights) or a scree slope (a gradual slope formed of loose rocks, which is safe to ascend but unsafe to descend). Utilities for each action, depending on the agent's type τ, are given in the table.

large benefit to those who can access it (i.e., swimmers), but incurs a small cost for those who cannot. Sampling only a few teachers may not result in observing any swimmers, so selection is highly sensitive to their frequency (and the cost against which it trades off). By contrast, sampling many teachers usually results in observing a swimmer; if any had a map containing the shortcut, they would capitalize on it and thus be selected. As a result, cultural evolution systematically leads to overly complex maps from a population-fitness perspective. Perhaps cultural evolution truly is a ratchet after all?

Note, however, that the foregoing results were driven by additional complexity giving some agents an edge (so to speak) over others. But another way that maps can confer differential benefits across a population is by protecting against *disadvantage*. We turn to this case next.

Protecting against disadvantage via a backup route

Suppose that, rather than planning a sedate journey from the village to the town market, agents must instead trek from the village along a harrowing trail that skirts a cliff edge to reach their goal, perhaps an important religious site. Now, unlike the market journey, where agents could always take the slower bridge route regardless of their swimming proficiency, the cliff edge poses an unavoidable threat to agents that are scared of heights. Acrophobic agents travel slowly along the cliff edge, earning less utility from the journey than do fearless agents. (We assume that agents' susceptibility to acrophobia is set at random, according to Pr(acrophobe).) Fortunately, there is an alternative "backup" route up a scree slope (a gradual ascent formed of loose stones). Using this backup route enables agents with fear of heights to travel as quickly as fearless agents. However, the scree is unsafe to descend, so all agents must return via the cliff edge trail. Figure 4 shows the minimal graph that illustrates this planning problem.

The key difference here is that, rather than offering an *advantage*, the backup route partially nullifies the *disadvantage* acrophobic agents face. Because the backup ascent is of equal length to the cliff trail, fearless agents do not gain utility by representing it. (Formally, this case is similar to the river journey, except that we have subtracted one util from the agents who would benefit from the non-standard route.) Will cultural evolution preserve this backup as it did the shortcut?

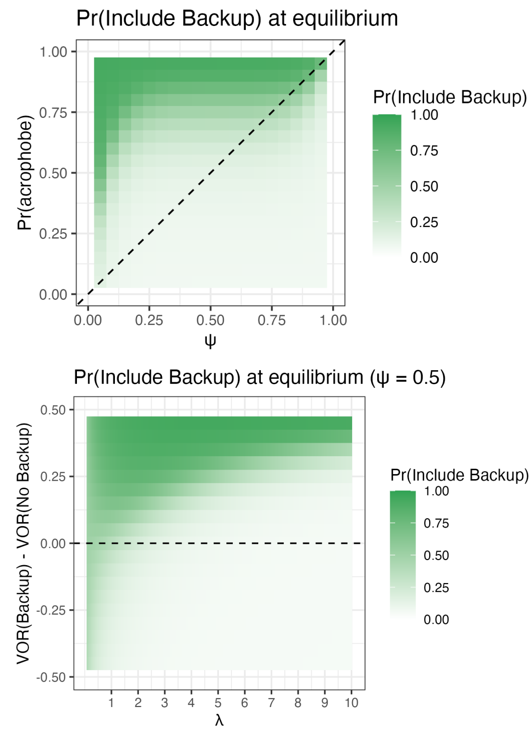


Figure 5: **Top:** The heatmap shows the probability the culturally evolved map includes the shortcut, as a function of Pr(acrophobe) and ψ (fixing $\lambda = 7$). **Bottom:** This heatmap shows the favored construal as a function of its VOR and the scarcity of social learning. Here, as social learning becomes more plentiful, cultural evolution fails to preserve the backup option even when doing so would maximize average fitness.

As before, the cost of complexity ψ and the probability that the backup route is helpful (i.e., Pr(acrophobe)) create a tradeoff that determines the cultural evolutionary equilibrium. But now, within this equilibrium, payoff-biased social learning exerts a bias toward hatcheting away the backup route, even when doing so decreases overall population fitness (Figure 5).

Examining both models together, evolutionary pressure toward simplicity when social information is scarce emerges from a general property of construals rather than details of the problem structure. Additional information is useful only to those who have the opportunity to act on it, but its associated complexity is costly to everyone. In the shortcut case, the highest-fitness agents are the lucky ones who were endowed with aquatic expertise and who acquired the shortcut map; they are copied even by agents who cannot themselves take advantage of the shortcut.

On the other hand, evolutionary pressure toward complexity is payoff-dependent. When a complex map protects against a *disadvantage*, cultural selection very rarely favors the complex map, because its use is never associated with especially "fit" behavior. That is why, in the backup case, the compensatory backup path is often not represented: Unlucky acrophobic agents can only match the fitness of a height-tolerant agent. When a complex map offers a distinctive *advantage*, however, best-of-k cultural selection often favors the complex map, because its use is associated with especially "fit" behavior.

Discussion

Here, we have presented a minimal model of culturally evolved task construals. By embedding a model of resource-rational representation into a cultural evolutionary process, we provided a proof of concept that cultural evolution may not always yield increasing complexity. Instead, when we account for the cost of representing complex cultural artifacts, we find that cultural evolution can sometimes push towards simplicity. Provided that more successful agents are more likely to pass down their construals, cultural evolution will favor construals that strike a balance between the utility obtainable when using that construal and the cost of representing it.

Crucially, however, cultures may not strike the balance that would be best for the population (i.e., that maximizes aggregate payoff). Instead, both the availability of social information and how agents use that information play a critical role in determining the construals a culture ultimately adopts. The sign and magnitude of this bias are governed by specific properties of individual variation and social learning, respectively. Culture is biased towards *complexity* when an innovation offers offers users a distinctive *advantage*, but is biased towards *simplicity* when an innovation protects users from a *disadvantage*. In both cases, the bias becomes increasingly strong as the number of teachers each agent has access to increases.

The specific way we modeled selective social learning (i.e., best-of-k), biased the cultural equilibrium away from the global optimum. Alternative approaches, such as assuming selection from the entire population proportional to fitness (Taylor & Jonker, 1978), would settle on the global optimum (because this rule implements an expected value calculation). Moreover, cultural artifacts are often produced with the express purpose of satisfying the diverse needs of a population. More generally, introducing pedagogical forms of transmission would likely push cultural evolution towards global optima (O'Madagain & Tomasello, 2021; Shafto et al., 2014).

Nevertheless, our findings concerning the cultural evolution of task construals echo certain trends in real-world cartography. We observed that maps evolve to strike an appropriate level of detail for the task at hand, as mediated through processes of cultural transmission. Likewise, maps produced for orienteering (which requires navigating rough terrain using a compass) often include topographic detail such as elevation isolines and the location of tree stands that casual trail maps (whose users follow a prespecified route) do not. Across different cities, and across time, subway signage has trended both in the direction of increasing abstraction (e.g., the London Tube post-1933) and toward increasing detail (e.g., the New York City MTA post-1972). We speculate that one driver of these cartographic trends may be the cognitive costs of planning.

Our model explores how planning costs can constrain the complexity of cultural forms. But other kinds of constraints likely matter too. In particular, models of language evolution often invoke a tradeoff between simplicity and informativity (Kirby et al., 2015), with informativity playing the role of utility in our model. In that context, simplicity is most often motivated by learnability (Carr et al., 2020; Griffiths & Kalish, 2007; Hofer et al., 2023; Kirby et al., 2008) or signal length (Piantadosi et al., 2011; Zaslavsky et al., 2018), although some accounts explicitly point to processing costs (Hawkins, 2004; Kemp & Regier, 2012). In our context, a complex map is harder to read (learnability), reproduce (signal length), and plan with (processing). Dissociating these three pressures will require precisely formalizing their differences, something that neither our model nor, to our knowledge, existing models of language have jointly captured within a single framework.

Task construals, and the cultural artifacts that instantiate them, are one example of causal models among many. Some theories of cumulative culture emphasize the importance of individually-learned causal models in explaining our species' success (Pinker, 2010). Other theories foreground the explanatory power of cultural variability, high-fidelity imitation, and cultural selection operating over beliefs and practices whose adaptive benefit need not be causally transparent (Boyd et al., 2011; Heyes, 2018). Our model illuminates a third way: the *interplay* between individual cognition (here, resource-rational planning) and cultural selection (via capacity-constrained social learning) plays an essential role in determining the outcome of cultural evolution.

Conclusion

Theories of cumulative culture tend to stress accumulating complexity, while theories of resource-rational planning often portray complexity as a challenge requiring simple approximations to surmount. Impelled by this tension, we brought these ideas together into a common computational framework and used it to model the cultural evolution of task construals.

By examining our model, we learned that cultures do not always ratchet useful complexity, nor do they strictly hatchet away burdensome detail. Instead, cultures calibrate the simplicity of socially transmitted artifacts and representations, balancing their utility against their costs of use. Moreover, the process of cultural transmission imposes its own biases, sometimes favoring a *miscalibrated* level of complexity.

These core insights suggest two promising directions. First, mirroring language evolution, our model predicts that experiments instantiating cultural evolutionary dynamics will regularize the complexity of culturally transmitted forms used in planning. And second, it may be that historical trends in material culture are driven in part by cognitive costs.

Acknowledgments

We thank Joe Henrich, Neil Bramley, Bonan Zhao, and members of the Laboratory for Social Cognitive Science for their feedback and guidance. XRG is supported by the Department of Defense through the National Defense Science and Engineering Graduate Fellowship Program. This work was supported by award number N00014-22-1-2205 from the Office of Naval Research (to FC).

References

Baldini, R. (2013). Two success-biased social learning strategies. *Theoretical Population Biology*, *86*, 43–49. https://doi.org/10.1016/j.tpb.2013.03.005

Barrett, B. J., McElreath, R. L., & Perry, S. E. (2017). Pay-off-biased social learning underlies the diffusion of novel extractive foraging traditions in a wild primate. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1856), 20170358. https://doi.org/10.1098/rspb.2017.0358

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, *108*, 10918–10925. https://doi.org/10.1073/pnas.1100290108

Callaway, F., Van Opheusden, B., Gul, S., Das, P., Krueger, P. M., Griffiths, T. L., & Lieder, F. (2022). Rational use of cognitive resources in human planning. *Nature Human Behaviour*, *6*(8), 1112–1125. https://doi.org/10.1038/s41562-022-01332-8

Carr, J. W., Smith, K., Culbertson, J., & Kirby, S. (2020). Simplicity and informativeness in semantic category systems. *Cognition*, *202*, 104289. https://doi.org/10.1016/j.cognition.2020.104289

Chater, N., & Christiansen, M. H. (2025). Social tinkering: The social foundations of cultural complexity. *Behavioral and Brain Sciences*, 1–62. https://doi.org/10.1017/S0140525X25103981

Enquist, M., Ghirlanda, S., Jarrick, A., & Wachtmeister, C. .-. (2008). Why does human culture increase exponentially? *Theoretical Population Biology*, *74*(1), 46–55. https://doi.org/10.1016/j.tpb.2008.04.007

Griffiths, T. L., & Kalish, M. L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, *31*(3), 441–480.

Hawkins, J. A. (2004). *Efficiency and Complexity in Grammars*. Oxford University Press.

Henrich, J. (2017, October 17). *The secret of our success*. Princeton University Press. Retrieved December 16, 2022, from https://press.princeton.edu/books/paperback/9780691178431/the-secret-of-our-success

Heyes, C. (2018). *Cognitive gadgets: The cultural evolution of thinking*. Harvard University Press. Retrieved October 26, 2022, from https://www.hup.harvard.edu/catalog.php?isbn=9780674980150

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2022). People construct simplified mental representations to plan [Publisher: Nature Publishing Group]. *Nature*, *606*(7912), 129–136. https://doi.org/10.1038/s41586-022-04743-9

Ho, M. K., Cohen, J. D., & Griffiths, T. L. (2023). Rational simplification and rigidity in human planning. *Psychological Science*, *34*(11), 1281–1292.

Hofer, M., Kirby, S., & Levy, R. (2023). Simplicity and informativeness in the evolution of combinatorial structure. *Proceedings of the annual meeting of the cognitive science society*, *45*(45).

Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, *336*(6084), 1049–1054.

Kendal, J., Giraldeau, L.-A., & Laland, K. (2009). The evolution of social learning rules: Payoff-biased and frequency-dependent biased transmission. *Journal of Theoretical Biology*, *260*(2), 210–219. https://doi.org/10.1016/j.jtbi.2009.05.029

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686.

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102.

Lieder, F., Callaway, F., & Griffiths, T. L. (2026). *The rational use of cognitive resources*. Princeton University Press.

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: Estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1509), 3515–3528. https://doi.org/10.1098/rstb.2008.0131

O'Madagain, C., & Tomasello, M. (2021). Shared intentionality, reason-giving and the evolution of human culture [Publisher: Royal Society]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *377*(1843), 20200320. https://doi.org/10.1098/rstb.2020.0320

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(9), 3526–3529.

Pinker, S. (2010). The cognitive niche: Coevolution of intelligence, sociality, and language [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *107*, 8993–8999. https://doi.org/10.1073/pnas.0914630107

Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology*, *71*, 55–89. https://doi.org/10.1016/j.cogpsych.2013.12.004

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

Taylor, P. D., & Jonker, L. B. (1978). Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, *40*(1), 145–156. https://doi.org/10.1016/0025-5564(78)90077-9

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1528), 2405–2415. https://doi.org/10.1098/rstb.2009.0052

Wu, C. M., Vélez, N., & Cushman, F. A. (2022). Representational exchange in human social learning: Balancing efficiency and flexibility. In C. M. Wu, E. Schulz, & I. Cogliati Dezza (Eds.), *The drive for knowledge: The science of human information seeking* (pp. 169–192). Cambridge University Press. https://doi.org/10.1017/9781009026949.009

Zaslavsky, N., Kemp, C., Regier, T., & Tishby, N. (2018). Efficient compression in color naming and its evolution. *Proceedings of the National Academy of Sciences*, *115*(31), 7937–7942.