Kim Sterelny’s book *The Pleistocene social contract* provides an exceptionally well-informed and credible narrative explanation of the origins of inequality and hierarchy. In this essay review, we reflect on the role of rational choice theory in Sterelny’s project, before turning to Sterelny’s reasons for doubting the importance of cultural group selection. In the final section, we compare Sterelny’s big picture with an alternative from David Wengrow and David Graeber.

**KEYWORDS**
cooperation, culture, hierarchy, human evolution, inequality

1 | THE RISE OF ELITES

*The Pleistocene social contract* is Kim Sterelny’s third book about human social and cognitive evolution. Although not intended as such, the books form a reasonably coherent trilogy—certainly more coherent than the last three *Star Wars* episodes. The first, *Thought in a hostile world* (Sterelny 2003), offered an account of the evolution of a form of belief-desire psychology—“decoupled representation”—that stressed both its distinctive value in early hominin social worlds and the distinctive value of being able to track it in others: theory-of-mind. The idea of a “debunking genealogy” of morality or religion is familiar. What Sterelny offered was an at least partially vindicatory genealogy of folk psychology.

The second episode was *The evolved apprentice* (Sterelny 2012a), a gradualist and multifactorial account of how humans evolved such a distinctive cognitive profile in comparison with our great ape relatives. Ideas sketched briefly in the first book were developed in much greater depth.

---

1The term “hominin” encompasses the *Homo* genus and the australopithecines; in other words, modern humans and all ancestral species following the split from *Pan*.
here. The idea of a “scaffolded” learning environment, where juveniles learn through structured trial-and-error in an environment richly seeded with the products of adult cooperation (everything from partially processed foods to partially made tools), is at the heart of the story. Hominins, in Sterelny’s picture, became increasingly reliant on this mode of cultural learning, driving life-history changes (a greatly extended childhood) and cognitive changes (adults evolved to be tolerant and supportive of their “apprentices,” who evolved to learn rapidly and reliably from their role models) that made them exceptionally good at it: evolved apprentices.

The third episode picks up where *The evolved apprentice* left off, recapitulating the main ideas before launching into a detailed account of how human social organization changed over time, with a particular focus on the Pleistocene–Holocene transition around 11,000–12,000 years ago. This is the transition in which humans across the world (and they were, by this point, modern humans, *Homo sapiens sapiens*) embraced farming and also, it seems, became far more tolerant of hierarchy and inequalities of wealth and status. The focus is primarily on social evolution—culture and cooperation—rather than cognitive evolution. A modern suite of cognitive abilities is assumed to be in place long before the Holocene, and the question is why, without underlying cognitive changes, social organization changed so dramatically. What Sterelny calls the “Pleistocene social contract,” in which aspiring aggrandizers were reliably knocked back by coalitions of groupmates, creating a tense kind of stability known as “forager egalitarianism,” dissolved in multiple independent events in different parts of the world. Why?

It was not the rise of the state. This puts the cart before the horse, because “transegalitarian” societies with marked hierarchies and inequalities precede well-organized, coercive states. Nor was it a case of the worst-off endorsing Rawls’s “difference principle” and choosing inequality for their own benefit. It is not credible that inequalities made the worst-off better off, since, as a tradition in anthropology starting with Sahlins (1968) has persuasively argued, farming is a grueling, miserable existence for those at the bottom of the emerging social pyramid, not at all to be preferred to foraging in environments with copious food. As Sterelny puts it, “foragers leave their excrement behind; farmers live amongst it” (Sterelny, 2021, p. 126).

So what happened? Sterelny’s account posits four “enabling conditions”: reliable food surpluses that can be stored, tension between neighboring communities, “clan-like” social organization, and widespread cultural transmission of ritual and religion. “Enabling conditions” make a transition possible without necessitating it: The idea is that the transition will happen only if other historically contingent factors are also in place, but no robust, general account can be given of these other factors.

Why were these four factors enabling conditions? Reliable, storable food surpluses create opportunities for aggrandizers who seek personal wealth to accrue it. Growing intercommunity tensions create opportunities for these aggrandizers to generate benefit for the masses by brokering peace with neighboring communities (e.g., by hosting large feasts) while at the same time ramping up the potential costs of within-group conflict aimed at bringing the aggrandizer down, since a disharmonious group will be more easily defeated. As Sterelny puts it, “the competitive and fraught relations between communities (and hence the advantage of signaling one’s wealth, strength and allies) opens this door to aggrandizers to turn wealth into political leadership, and leadership into wealth” (Sterelny, 2021, p. 139). Clan-like organization, in which patrilineal (or, in some cases, matrilineal) kin groups have strong obligations to assist each other despite geographic separation, creates networks that “enable would-be aggrandizers to recruit social and material support, and through which that support could be rewarded” (p. 151).

Finally, and most provocatively, the widespread cultural transmission of ritual creates a different kind of elite: An epistemic elite regarded as possessing esoteric knowledge of the supernatural
and trusted with leading group-wide rituals. This enables mutually beneficial alliances to form between the epistemic and material elites: Shamans, in return for material privileges of their own, purport to reveal hidden truths favorable to the continuation of the unequal social order. “The ritual and religious lives of these transegalitarian communities became channels through which the community is half persuaded, half coerced into accepting the new status quo” (p. 145). Religion, on this account, is the handmaiden of inequality, hierarchy, and living in shit.

These transegalitarian communities were what Dennett might call hemi–demi–semi states. There was no effective coercive apparatus, but the real threat of war with the neighbors, along with the bogus but credible threat of supernatural consequences, could be dangled before the non-elite masses as a form of quasi-coercion, a sufficient motivation to keep the tribute flowing. Meanwhile, opting out—returning to a forager existence—became ever less viable as settled communities engulfed the most favorable foraging environments, pushing foraging communities to where they remain to this day: the margins.

This is “big picture” narrative explanation—but of a distinctive kind. Big pictures are often encountered in commercial books, where it is important to entertain. If attention to detail loses out in the trade-off, so be it. Speculative stories of prehistory are sometimes also used to make philosophical points, as in the work of writers as far-removed from each other as Thomas Hobbes and Philip Kitcher (2011). For Sterelny, by contrast, narrative explanation is the serious business of constructing an account that aims for as much robustness and generality as can be achieved, and that joins up with archeological, ethnographic, psychological, and primatological evidence in as many places as possible, as well as with the results of mathematical modeling and simulation. The result is an exceptionally well-informed and credible account.

Nonetheless, our aim in the following sections is to find a few points where the account can be challenged. We will begin with some reflections on what can (or cannot) be safely assumed about the rationality of early humans, before turning to the significance (or not) of cultural group selection. In the final section, we compare Sterelny’s big picture with an alternative (or is it?) from Graeber and Wengrow (2021).

2 | THE RATIONALITY OF EARLY HUMANS

To explain transitions in social organization, we need to be willing to impute at least some rationality to early humans. We need to make some assumptions about their goals, and we need to assume a certain level of stability in these goals across geographic regions and across generations. Yet we also need to make room for ways in which early humans were (like us) far from ideally rational, and to allow for the possibility of both synchronic variety and diachronic change in their goals and values. How do we strike that balance? This is one of the major methodological choices that confronts anyone attempting to build an account of human social evolution.

Sterelny’s account provides an interesting case-study. As his attention turns from the Pleistocene to the Pleistocene/Holocene transition, he becomes notably more willing to employ the concepts of “rational” and “irrational.” This is in line with his qualified endorsement, in past work, of a rational choice framework for understanding prehistory (Sterelny, 2012b, 2015). His explanation of the origin of farming, for example, is an explanation of why it would have been rational for a fitness-maximizing agent to adopt farming, even granting the Sahlins view that the transition involved a gradual decline in the quality of life of the worst-off. It is an account of how incremental adjustments that would have seemed “minor and sensible” (Sterelny, 2021, p. 132) at the
time might nonetheless have led to an end-state worse than the now-irrecoverable initial state. An important idea in the background is the tragedy of the commons (Hardin, 1968). A population of humans making individually rational choices to dial up the intensity of their storage and management of food resources may nonetheless find themselves depleting the environment to such an extent that they foreclose the option of ever returning to foraging in a world of abundant food.

This rational choice framework leads to some deep puzzles. Consider, for example, Sterelny’s brief account of the origin of property rights:

Farming also encourages a shift to norms that respect property rights and to formal or informal sanctions for violating them (Bowles and Choi 2013; Gintis 2013). Farmers invest in their farms and crops: initially, by holding back seed to plant, and then clearing land, in preparing and improving soils; in tending crops; in investing in tools to plant, harvest, process, and store produce. They invest in a built environment, not just in their lands and crops. These investments in crops and buildings would be profoundly irrational without secure possession of the product. Property rights function as guarantees of secure possession. (Sterelny 2021, p. 129).

The puzzle here is that of how, prior to the existence of states, there could have been “guarantees” of security that were solid enough to make investment in land rational for an individual seeking to maximize fitness. There is no state to enforce them, and the informal coalitions of “forager egalitarianism” are not around to enforce them either. Who is? In the Bowles and Choi (2013) model Sterelny cites, farming and strong social norms of respect for property rights proliferate as a package deal under group selection when they emerge together by chance—but this happens in only 31 of 1000 simulation runs. The mystery only deepens when we note that, to really make investment rational for a fitness-maximizing agent, rights of inheritance also have to be respected, as Sterelny points out. This makes the “fluke simultaneous emergence” hypothesis even less plausible.

Gintis (2007, 2013) has a different approach to property rights, one that rests on an analogy with territoriality. In many animals, conflicts over territory are often resolved non-violently in favor of the incumbent. This is plausibly because violent conflict is extremely costly, and a default policy of resolving conflict in favor of incumbents is an evolutionary stable strategy for avoiding those costs. But as Gintis notes, there is also an “anti-property equilibrium” in his model, where conflicts are resolved by default in favor of migrants. Gintis appeals to group selection as an explanation for why real equilibria tend to favor incumbents. But Sterelny is generally skeptical of appeals to group selection in the human case, for reasons we will discuss in the next section. So, the chicken-and-egg problem regarding norms of property ownership and investment in land remains stubbornly in place.

Elsewhere in the account, meanwhile, departures from the rational pursuit of fitness play a crucial role. In earlier work, Sterelny (2012b) described the transition to transegalitarian societies as involving a decoupling of fitness and utility. Pleistocene humans, although cognitively capable of pursuing goals that undermined their biological fitness, are unlikely to have done so. Cultural inheritance would have been mostly vertical (i.e., from biological parent to biological offspring), and in such a situation the cultural variants most likely to spread are those that align well with biological fitness (Birch, 2017; Birch & Heyes, 2021). In the Holocene, by contrast, cultural learning runs amok, with the parent-offspring channel an increasingly minor part
of it, and cultural variants that detract from biological fitness can rip through the population—going viral—if they transmit readily enough. Religious doctrines that justify inequality are posited by Sterelny to have something like this character. They spread because of the prestige of those who promulgate them (shamans and chiefs), even though they are “not in the best interests of those who take them up” (Sterelny, 2021, p. 144). So, in this particular domain, early humans are posited to be rather poor judges of their fitness interests. This is compatible with them still being rational, of course, because we can think of it as a change to the goal they are rationally pursuing. The goals start to include access to an afterlife, avoidance of divine punishment, and so on. Human goals and values are increasingly untethered from both fitness and reality.

This leads to a methodological worry about the rules of the game. What are the constraints on postulating changes to the utility functions of early humans, such that they come to pursue goals that oppose their fitness interests? In principle, one could make all puzzles about human social evolution dissolve this way. Property rights? Well, maybe agents just came to value property rights for their own sake, under instruction from the local elites who benefited most from those rights. Yet this feels illicitly ad hoc—it feels like cheating. But then, whatever the source of this intuition of illicitness, surely it is also present when the question is why people acquiesced to elites in the first place.

There is a deep issue here, an issue about how many “get out of fitness-maximization free” cards one has, and how they can be played. The decoupling of fitness from utility creates a situation where the old rules of adaptationist explanation are being dissolved by the evolutionary process itself, yet the modern rules of economic explanation are not yet fully applicable. In this liminal zone, it is not clear what the rules of explanation are.

3 IS CULTURAL GROUP SELECTION TOO QUICKLY DISMISSED?

At several points in the book, Sterelny downplays the importance of cultural group selection (CGS), in which competition between groups drives cultural change (Richerson and Boyd, 2005; Smith, 2020). He is skeptical that CGS was a frequent or important force shaping forager life in the Pleistocene. The foragers of the Pleistocene—from *Homo heidelbergensis* onwards—had sufficient individual-level incentives for cooperation between bands, for avoiding conflict between bands, and for de-escalation of conflicts that did arise (Sterelny, 2021, sections 3.1 and 3.3). And there is little archeological evidence that would support rampant conflict among populations—at least not until the end of the Pleistocene and the Pleistocene–Holocene boundary (Sterelny, 2021, pp. 115–117).

Sterelny attributes the greater importance of between-group conflict around the start of the Holocene to the rise of storage, sedentary settlements, property norms, and the accumulation of social capital. These ecological circumstances can feed intergroup conflict because stored food (and later, symbolically valued goods) and resource-rich ecologies (such as maintained fields) offer a valuable target for other groups. Moreover, groups around this time were often (though not always) organized into segmented lineages and clans; the kinds of corporate entities that can mobilize an armed force to raid or seize territories and goods (Sterelny 2021, section 4.4).

Yet even in situations that might have favored between-group conflict, Sterelny remains skeptical of the role of CGS. He gives two reasons for skepticism: First, norms, practices, and
institutions are not “atomistic” traits and thus are poor targets for selection; second, the population of groups was too small for CGS to be a powerful force.

Let us first consider the claim about “atomism.” The idea is that group-level practices (norms, institutions) are too integrated with each other, too interdependent, for selection to be able to shape them efficiently. Integrated sets of group-level practices face selection as a whole, with no room to tweak and improve any particular component part independently of the rest. It is a familiar idea in evolutionary biology that different phenotypic characters have to be “quasi-independent” (Lewontin, 1978) for selection to tune adaptations effectively. “Quasi-independence” means that different characters, though not fully independent, can vary and be selected independently from each other. Sterelny (1992) once argued that, for this reason, individual behaviors should not be regarded as adaptations. Selection cannot shape a single behavior without altering the underlying psychological mechanisms that produce it, thereby altering many other behaviors. So it is the psychological mechanisms, not the behaviors, that are the adaptations. Thirty years later, he is directing this complaint at social norms and institutions.

But is it true that social norms and institutions fail the test of quasi-independence? Let us approach this issue indirectly, by first considering individual-level cultural products. A bow is in some ways a highly integrated object, in that tweaking one aspect (like the length of the string) will tend to undermine the functionality of the whole item unless there are compensating changes elsewhere (to the length of the bow). But cumulative cultural evolution of bows is still possible. This because the structure of the object is transparent to intelligent agents (Brown, 2021). It does not take a deep understanding of the structure of the bow to realize that modifying one component necessitates compensating modifications to other components. So, we can see how a process of cultural evolution might cumulatively improve a bow, even though a process relying on entirely random changes to individual components might not get very far. Variation is guided just a little bit towards sensible options, and this can compensate for strong interdependence of different components. This suggests that cultural evolution may not always need the same level of quasi-independence as its biological analogue.

Might group-level traits at least sometimes be as transparent as bows? Consider rituals. Elite individuals, such as shamans, might well be aware of the functions, components, costs and benefits of rituals they have led or witnessed multiple times, such as rituals of group admission (to adulthood, male cults, etc.), or of social cohesion (dances, feasts). They are also well placed to modify these rituals in sensible ways that do not destroy their functionality. They might modify myths or narratives to amplify ritual significance, or adjust the frequency of such rituals, altering levels of group allegiance and commitment to shared projects. Atran (2016) has observed this kind of “intelligent” modification of rituals and myths in the practices underwriting contemporary terrorist groups.

Religion seems a good target for CGS. Recent work has stressed religion’s role as a flexible, ecologically sensitive, and adaptive evolutionary system, with variety in ritual costs, context, timing and intensity as well as the domains, knowledge, and intentions of supernatural beings, varying with socio-ecological circumstance (Purzycki & Sosis, 2022). There is also ethnographic evidence suggesting that periods of ritual innovation and invention occur during periods of strife and conflict (Whitehouse, 1995). So, the ingredients are in place for religion to be, as Purzycki and Sosis (2022) describe it, “an adaptive system where each of its constituent parts comprises an ever-changing, adaptive amalgam of concepts and behaviors” (p. 107). This is a picture of religion as a medium in which pro-social practices, conventions, and norms can be tweaked and optimized, without uncontrolled knock-on effects on other components. Different
Doctrines and rituals can be adjusted independently of each other, perhaps with some compensating adjustments.

These are reasons to think group-level traits may be easier to be fine-tune, sufficiently independently of each other, and more likely to be cumulatively improved through CGS, than Sterelny suggests. But what about the second reason for skepticism: population size? The thought here is that populations of groups around this time were too small to support powerful CGS. Population size matters for multiple reasons. One is that the smaller the population, the less likely it is that beneficial changes to existing structures will arise. Although detailed modeling would be needed to seriously investigate this concern, it may be slightly mitigated by our comments above. For if the structure of the relevant norms and institutions is reasonably transparent, and if elite individuals have substantial power to make small adjustments, adaptive variations may arise a lot more frequently than they would in biological evolution, potentially allowing CGS to generate cumulative adaptation in smaller populations.

4 | SEASONAL AND PERMANENT INEQUALITY

Sterelny’s book has landed at a moment of unusually heated discussion about the origins of inequality, a debate ignited by Graeber and Wengrow’s controversial book *The dawn of everything* (2021). Graeber and Wengrow’s central thesis is that there was no “origin of inequality” or “transition to inequality” as such. Rather, humans in prehistory were continuously trying out different forms of social organization at different times and in different places, with some temporary forms of inequality and hierarchy existing at least as far back as 40,000 years ago. Graeber and Wengrow argue for a reconfiguring of the traditional puzzle: It is not one of how inequality arose, but one of how we got stuck in one particular highly unequal form of social life, when human communities had previously been able to move fluidly between different forms of political organization.

It would be fascinating to hear Sterelny’s reaction to that book. We naturally wondered: Are the differences between Sterelny’s big picture and Graeber and Wengrow’s alternative just differences of emphasis? Or is something deeper at stake?

Here is one possible point of disagreement. Sterelny stresses how ecological factors contributed, alongside domestication and storage technology, to a reasonably robust transition from forager ways of life to transegalitarian societies. The transition passes through an intermediate stage: storage foraging. We mentioned this earlier as one of the “enabling conditions” for the evolution of inequality. Storage foraging is likely when there is predictable and large seasonal variation in resource availability, abundant resources in the high season, a seasonal spike that can be harvested efficiently, and a seasonal surplus that can be stored with low wastage and little risk (Sterelny, 2021, p. 127).

Graeber and Wengrow (2021), meanwhile, give various examples from ethnography in which seasonal surpluses lead to seasonal, not permanent, inequality or hierarchy. Their examples include the Inuit, the Kwakwaka’wakw (Kwakiutl) of the Pacific Northwest, and Plains Nations such as the Lakota and Crow. These populations experience seasonal abundance in foodstuffs, such as the large salmon-runs of the Pacific Northwest and the seasonal migration of buffalo through the interior of what is now the United States and Canada. These resources are harvested and surpluses are stored. For instance, there is well-documented evidence of the food storage techniques of the Pacific Northwest nations who relied on a broad spectrum of resources preserved through fermentation, drying, smoking, and salting. And similar
techniques supported buffalo-hunting nations—like the Lakota—who stored various game meats through drying, smoking, and preservation in fat.

What we see in these cases, say Graeber and Wengrow, are seasonal inequalities and hierarchies. For example, during buffalo runs, the otherwise “small and mobile bands” (Wengrow & Graeber, 2015, p. 607) of Lakota adopted “unequivocal authoritarian” tactics, appointing a group of warriors to act as police. These warriors would enforce norms related to the buffalo drive—punishing those who hunted clandestinely—and beyond, suppressing war parties, directing mass movement, and “otherwise maintaining law and order” (Lowie, 1948, p. 18). After the hunt, and especially during the winter months, the Lakota would return to smaller units that lacked such a disciplinary organization.

More generally, Graeber and Wengrow argue that many elements of social organization can vary seasonally. The precise changes in organizational form might differ from population to population, but extend beyond shifts in population size and the nature of authority. Populations may have different laws, rituals, social protocols, and even individual names from one season to the next. Graeber and Wengrow write that “archaeological evidence is piling up to suggest that in the highly seasonal environments of the last Ice Age, our remote ancestors were behaving much like the Inuit, Nambikwara or Crow” (Graeber & Wengrow, 2021, p. 111). They look at some of the same examples discussed by Sterelny, such as Gøbekli Tepe in Turkey, and argue that they point towards seasonal rather than permanent inequality.

Graeber and Wengrow’s broader argument is that ecological and demographic conditions are not, in ethnography or prehistory, determinative of political organization. We should instead recognize our ancestors as having deliberately experimented with different subsistence and organizational forms, moving fluidly between different options. Though organizational form might vary with the seasons, this did not mean that the ecology determined ritual, governance, or social protocols. As they see it, the transition to sedentary agriculture, larger cities, and large (but efficiently managed) population sizes was far from robust.

The case for reading this as a difference of emphasis is that Sterelny only regards ecological and demographic factors as “enabling conditions” for the evolution of inequality. They make it possible; they do not guarantee it. Graeber and Wengrow might agree. Moreover, one need not see seasonal variation in organizational form as being in tension with the broader ecological story told by Sterelny, if the seasonal variation has ecological drivers. But there is at least the hint of a deeper disagreement too, in so far as Sterelny does not regard it as a doomed project to look for generalizable and robust mechanisms for generating and stabilizing inequality and hierarchy—mechanisms not specific to any particular time or place—and he clearly thinks ecological and demographic factors will be an important part of any such mechanism. Graeber and Wengrow might disagree.

This relates back to one of our earlier reflections. One response to these examples of seasonal inequality is to postulate that norms of land ownership, and land inheritance, may be even more central to the stabilization of permanent inequality than Sterelny’s account suggests. It is not storage foraging by itself that sets the transition in motion. It is really the idea of ownership—ownership that persists through the seasons and across generations—that allows inequality to bed in permanently. And yet, as we pointed out earlier, Sterelny’s picture, and the modeling work he cites, leaves the origin of ownership norms rather mysterious. The mystery can be partially though not wholly diminished if we are willing to posit a form of cultural group selection powerful enough to favor both investment in land and social norms that support claims of ownership.
DATA AVAILABILITY STATEMENT

There is no data.

ORCID

Jonathan Birch https://orcid.org/0000-0001-7517-4759
Andrew Buskell https://orcid.org/0000-0001-6939-2848

REFERENCES
