From the existential stance to social constraints

How the human mind becomes embedded in our social, cultural and material context

Avel Guénin—Carlut

University of Sussex, Active Inference Institute, Kairos Research

avel@kairos-research.org

Abstract

Contemporary social and cognitive science rely on the methodological separation between cognitive architecture on one hand, and the content of cultural norms and attitudes they integrate throughout their life on the other. Yet, a careful consideration of human development and of the deep history of human minds call into question the validity of this separation. We propose here a physical ontology describing the nature of social / cognitive structure, as well as as its relation to agency. Based on Active Inference, a computational and mathematical theory describing human cognition as a permanent prediction of one's sensori-motor flow, we conceptualize the novel notion of *social constraints*. Social constraints formalizes the embedding of agent's cognition in a shared social context, if those agents are prone to reflexively consider their own beliefs and actions as communicative acts in a cognitive attitude we call the *existential stance*. By analogy to gauge theory, the main foundational theory of contemporary physics, social constraints exist dually as formal symmetries over a system (here, prior beliefs over the causes of the agent's sensation and actions) and as causal forces driving its dynamics (here, constraints over the agent's cognitive activity, especially regarding the flow of their attention). But unlike the physical theories describing inert matter, social constraints are embedded in a shared material, social, and cultural environment which encultured agents must actively integrate and (re)produce through their activity. We consider the implications of this active process of integration and (re)production of social constraints, both for the underlying physical ontology and for the methodology of social and cognitive sciences. In particular, we provide a detailed account of the relation between social structure and cognitive agency, and we ground a description of social organization as *bona fide* biological organization embedded in the cultural and material context for human activity. Overall, our argument provides what is to our knowledge the first fully coherent physical ontology for the development of human cognition and the evolution of sociocultural forms of life.

Table of Contents

0) Introduction

[0.1] The idea of a science of society modeled after physics, with its well defined laws inferred from hypothetico-deductive reasoning, has been a major motivation of sociology since the foundational work of Auguste Comte (Comte 1844). In contemporary terms, Comte's positivism was best understood as a philosophical and religious *credo* for reason and progress, rather than a strictly epistemological or scientific program. Yet, it exerted an immense influence over the development of the scientific methodology throughout the later XIXth and XXth century. Most notably, positivism served as a blueprint for the methodology of sociology through Durkheim's « Les Règles de la méthode sociologique » (Durkheim 1895 (1982)), and for the entirety of contemporary philosophy of science through Neurath's and colleagues' formalization of logical (or empirical) positivism as an epistemological framework. The strictly positivist conception of scientific methodology did not hold water very long. Indeed, it suffered from the devastating argument by Duhem (Duhem 1906 (2016)) then Quine (Quine 1951) that single propositions cannot be verified or falsified in isolation from a broader formal framework. However, positivism did survive through its reinvention by Durkheim, who interpreted the formal relations between statistical quantities as the outcome of underlying social structures. The power of structural explanation came from the fact that it mimicked the formal regularity of lawful explanation, while remaining coherent with the variety of behavior that *a priori* similar systems could display. In its most classical form (*e.g.* Lévi-Strauss, 1958), structuralism would consider linguistic or social activity as a simple instantiation of prior social or semiotic systems, therefore redefining the object of science as "social structure" rather than "laws of history" while keeping a similar strategy for scientific inquiry as positivism.

[0.2] In the latest half of the XXth century, the life sciences progressively the standard model of scientific methodology, and physics was reformulated by calling onto underlying structure through what is known today as "gauge theories". Hence, the ability of scientific theory or models to capture causal structure in the world became the gold standard of scientific inquiry. This is especially straightforward in the life sciences and cognitive neuroscience, which served as the basis for the mechanistic model of scientific explanation (Bechtel and Abrahamsen 2005; Bechtel 2008; Boone and Piccinini 2016). However, the specific understanding of what causal structure is relevant to explaining human behavior remains a controversial topic in the human sciences. At one extreme, proponents of Evolutionary Psychology (Barkow, Cosmides, and Tooby 1995; Cosmides and Tooby 2007; 1992) and the associated Cultural Attraction Theory (Claidière and Sperber 2007; Sperber 1996; 2001; 2011; Sperber and Hirschfeld 2004) hold that both the universal and particular aspects of human behavior need to be explained by calling onto a supposedly universal and genetically encoded cognitive architecture implemented in the brain. At another, post-structural theorist such as Foucault (1980), Baudrillard (1994) or Latour (2005) (or, although too early to be qualified as post-structuralist, Goffman 2022 (1956)) would typically place the specific social, cultural and discursive context in which a specific agent acts front and center of any explanation of behavior. By extension, both schools rely on wildly different methodologies to explain human behavior – one extrapolates cognitive universals from contemporary behavioral regularities as measured in experimental conditions, the other draws massively from qualitative analysis of idiosyncratic patterns embedded in locally afferent narratives or social relations.

[0.3] While often decried by proponents of methodological naturalism, this diversity in the targets and the methodology of various approaches in the human has nothing surprising. Indeed, for all the conceptual coherence it brings to the table, organizing scientific inquiry around the discovery and description of causal structure as it exists in the world brings novel challenges which are not shared by law-based approaches. A scientific law is by construction a prior fact of nature, which need and may not be described beyond its manifestation. On the other hand, a mechanism is a specific thing that exists at a specific place, and whose observable behavior is the result of a specific material structure as well as the specific context of its activity (Bechtel 2009). Even after being identified, a mechanism needs be localized, analyzed into individual components, their organization reconstructed, and their context of activity assessed before getting anywhere near a lawful input-output law. Moreover, the explanation of natural phenomenon by causal structure is not self-standing in the same manner that law-based explanations are. Indeed, causal structures are actual natural objects, whose existence and organization are themselves a natural phenomena subject to explanation. These considerations strongly suggest that pluralism is a necessary trait of any science studying complex systems, as developed in Mitchell (2003). Perhaps most importantly, they seem to suggest that an integrated study of the human mind and social activity is an impossible and vacuous project, resulting only in an infinite regression of explanations to ground even mundane observation about social life. The cognitive mechanisms by which human integrate culture and the idiosyncratic narratives that dominate a given social context may constitute equally valid, but ultimately separate domains of scientific inquiry (as suggested *e.g.* in Ogien (2011)).

[0.4] The problem with this view appears clearly when we consider the diachronic evolution of the human mind and social organization in developmental or phylogenetic time. When considering the development of a human child, or when considering the evolution of the human mind, we do not see two distinct stages of successive development / evolution of "cognition" and then "culture". The development of a mature human mind depends critically on the existence of an adequate material and social niche, that supports the young during the many years they need to become able to support themselves and provides an adequate environment for the social learning of the adequate know-how to do so (Boyd, Richerson, and Henrich 2011). For example, the control of fire for food processing appears to be a critical adaptation for the development of our metabolically costly brains and endurance running abilities, dating at the very latest from the emergence of the *Homo sapiens* and *Homo neanderthalis* lineages and perhaps as far back as the emergence of *Homo erectus* (Wrangham 2017). Hence, the development of cognitively mature human adults relies critically on the existence and maintenance of the cultural technology that fire use constitutes. Furthermore, and although we may not rely on direct archaeological evidence for that, it seems straightforward that developing the basic social learning skills and practices to enables cultural transmission of complex technologies would have myriad consequences on the development of the human mind. Indeed, they enable the evolution of "cognitive gadgets", culturally transmitted cognitive mechanisms which psychologist and evolutionary anthropologist Cecila Heyes argue feed back onto the very mechanisms of social learning (Heyes 2018). This thesis, if verified, would entail the dissolution of a methodological distinction between "cognition" and "culture" as distinct (if related) areas of inquiry.

[0.5] Hence, we suggest that a principled understanding of human cognition requires an adequate account of culture and social context as a fundamental element of our mental life. The core difficulty in doing so is that, unlike a supposedly ancestral cognitive architecture, culture and social context change with each generation – if not every single instance of social interactions. Critically, these changes cannot, for the most part, be explained by outside events. An invading force may bring down an existing

State, a flood may destroy entire regions and their inhabitants, or an exceptional astronomical event may inspire fear and awe. But ultimately, how human communities make sense of such events, and the way they integrate them in their social practices, is up to their own agency. In other words, the processes by which human culture is transmitted are reconstructive in nature. There does not exist a physically plausible mechanism by which individual humans may directly copy any quantity cultural information embedded in the mind of their peers. Instead, they may infer what is the social and cultural context in which they engage in a robust but reconstructive process, which has the potential to drive cultural stability as well as open-ended innovation. Here, we propose a principled account of how the human mind is embedded in its social and cultural context which integrates at its core the way cognitive engagement participates in the evolution of this context. Our approach will focus first and foremost on the physics of cognition, and how they can provide an integrative model of the constructive relation between human mind and culture. Furthermore, we will discuss the limitation of a physicalist approach to human sociocultural evolution, and the way these limitations may constrain and inform further inquiry.

[0.6] To do so, we will rely mainly on two closely related framework aiming to describe and explain human cognition: the Active Inference Framework (AIF) and the Free Energy Principle (FEP). Both stem from the initial argument by Karl Friston and colleagues that the human brain, to produce efficient inferences about the causes of sensory stimuli and the adequate manner to respond to it, must minimize an informational Variational Free Energy (VFE) (Friston, Kilner, and Harrison 2006; Friston, Daunizeau, and Kiebel 2009; Friston 2010). Indeed, in the context of statistical inference, VFE minimization enables near-optimal inference in a very efficient manner, assuming certain mathematical constraints on the space of solutions. This literature rapidly developed in two complementary directions, respectively called the “low road” and “high road” to Active Inference by Parr, Pezzulo, and Friston (2022). The AIF, associated to the “low road”, aims to explain the mechanism of efficient, adaptive cognition by agents such as ourselves. It posits that agents understand and navigate their world through continuous prediction of their sensori-motor activity, which coherence with their prior models is maintained by VFE minimization in both their perception and action (Friston and Kiebel 2009; Friston, FitzGerald, Rigoli, Schwartenbeck, O’Doherty, et al. 2016; Friston, FitzGerald, Rigoli, Schwartenbeck, and Pezzulo 2016). The FEP, associated to the “high road”, starts from the question of how complex systems may maintain their own organization in precarious circumstances. This led to the development of an increasingly elaborate mathematical framework, demonstrating that existence (in the specific sense of individuation from one's world through a statistical boundary called a Markov Blanket) is dual to the minimization of VFE (Friston 2012; 2013; 2019; Da Costa et al. 2021; M. J. Ramstead et al. 2022; Friston et al. 2023).

[0.7] Together, those two lines of argumentation entail that, in a fundamental sense, any existing system enact a model of its world, and that the phenomenon of cognition reduces to a complex instance of this self-affirming process. This applies straightforwardly to describing the phenomenon of encultured cognition as the integration and enaction of world-models embedded in social relations and material culture, as exposed in (M. Ramstead, Veissière, and Kirmayer 2016; Veissière et al. 2020). In this context, the social and material context act to constrain the "regimes of attention" deployed by the encultured agent, which is brought to act through norms that are in a meaningful sense embedded in their environment (Guénin--Carlut and Albarracin 2024). However, the diachronic nature of cognitive evolution remains missing from those account. Indeed, the mathematical representation of "regimes of attention" in the FEP only explains attraction toward an optimal statistical model of external objects, and it does so by assuming the identity between the space of objects that actually exist in the external world and the space of perception by the agent. As a result, it cannot describe the construction of novel units of perception or action that agents may use to regulate their behavior, nor can it explain the active re-definition of interaction norms in a fixed environment (Guénin--Carlut 2023; 2024). Hence, the reconstructive aspect of enculturation under the Active Inference limits to deviation from optimality in cultural transmission, and may not drive the open-ended evolution of cognitive capabilities through human culture. Here, we propose to extend this account through a framework centered on the integration and (re)construction of "social constraints", understood as forces or structures embedded in a given social and material context and reconstructed through the active engagement of encultured agents.

[0.8] In the first section below, we propose to articulate a narrative theory of human cognitive evolution based on existing work studying the role of culture in comparative cognition and evolutionary anthropology. Our core argument is that the evolution of human cognition relies critically on the capacity to understand one's own actions reflexively in the context of one's social environment, a practice we call the *existential stance*. We argue that the existential stance enables the integration of one's self-model in abstract social expectations or semiotics, and therefore the evolution of complex social and cognition regulation. In the second section, we explain for the Active Inference model of enculturation, and we expose how the evolution of the existential stance in Active Inference agents would ground to a great extent the narrative model articulated in first section. In the third section, we dive deeper into the physics of cognition by exposing how encultured cognition can be described as the product of "gauge force", which is a type of fundamental force emerging from the implicit geometry of a system (here corresponding to cultural expectations embedded in a given material and social environment). Furthermore, we introduce the idea that those forces are best understood as "social constraints", a novel theoretical object which exist dually as an abstract expectation embedded in a given social context and as a concrete force acting on encultured agents' attention. In the fourth section, we expose the limitation of our account in accounting for open-ended evolution. We argue that open-ended evolution is a product of the contextuality of social constraints, *i.e.* the fact they exist only through their perception by given agents. Finally, in the fifth section, we outline that the present work opens a novel account of human sociality as the product of sociocultural forms of life (strictly speaking) embedded in the social and material niche of encultured humans. Furthermore, we outline the various challenges and opportunities entailed by such an account.

1) The existential stance and the cultural evolution of social regulation

[1.1] Human ecology is characterized by an extreme and exceptional reliance on "culture" (Boyd, Richerson, and Henrich 2011; K. N. Laland, Odling-Smee, and Feldman 2001), a somewhat loose term referencing the social transmission of information and its content. Indeed, no other animal species has relied on the widespread use of culturally evolved technologies like advanced gathering or hunting strategies, fire, clothing, and advanced tool fabrication to colonize the near-totality of Earth's earth surface area. Conversely, no other species would plausibly go extinct if even one generation lacked cultural learning, mainly through adequate tutoring by their elders. However, the reasons underlying the evolution of culture itself remains to this day a puzzle. By Heyes (2020)'s notation, the use of the term "culture" can range from Culture-1 (the simple social transmission of information) to Culture-3 (full-fledged Darwinian evolution based on cultural inheritance). While Culture-1 is nearly universal in animals, Culture-3 is relatively rare. Yet, Culture-3 has been shown experimentally to exist in non-human animals (Sasaki and Biro 2017), and it also seems to fit the widespread ecological pattern of tool use in apes and corvids. At a first glance, this seems to negate the view that there exists any unique human cognitive ability which enables full-fledged cultural evolution, and explains our unique cultural ecology. Indeed, developing empirical research has increasingly supported the view that the human evolutionary complex is not explained by a Big Special adaptation (borrowing here again Heyes' terminology in Heyes (2018)) but by the mosaic evolution of Small Ordinary traits working together to define the unique aspect of human ecology. Critically, those traits may be the result of genetic evolution, but they may as well be the result themselves of cultural adaptation (Heyes 2012a; 2018), or they may become embedded for reasons unrelated to direct selection in the material and social niche which human so carefully construct (Sterelny 2007; 2011; Boyd, Richerson, and Henrich 2011).

[1.2] The comparative study of human cognition with this of other animals should give a straightforward argument that no single traits constitutes a self-standing explanation to human specificity. Such a trait would by definition be necessary and sufficient for the development of human-like encultured cognition, and therefore universal in all human ecologies as well as exclusive to humans. Furthermore, it should itself be well-defined enough to provide a coherent mechanical or processual explanation of the phenomena of enculturation (see Bateson and Laland (2013) on evolutionary explanation). The most obvious candidates are cognitive abilities that are directly relevant to social learning, namely imitation, mentalizing and mirror neurons. Other candidates are involved in abduction more generally, namely causal reasoning and metacognition. Finally, other candidates are not causally effective in learning, but enable a learning-heavy ecology, namely low aggressivity, cooking, grand-mothering, long rapid eye movement (or paradoxical) sleep, big brains and long infancies. All were shown to be insufficient to explain the specificity of human cultural evolution. For example, relatively solitary corvids show human-like levels of mentalizing and causal reasoning while not displaying cumulative cultural evolution in ecological settings (Pika et al. 2020). On the contrary, pigeons with low cognitive abilities can show cumulative culture in specific experimental settings (Sasaki and Biro 2017). Finally, Heyes' work has pointed out that many candidates constitute ill-specified explanations, standing either as a correlate or byproduct of cultural learning. For example, mirror neurons constitute a simple homunculus of the capacity for social learning (Heyes and Catmur 2022), and so does imitation (Heyes 2012a). The practice of social learning seems to do a better work of explaining enculturation. However, it constitutes a socially contingent activity rather than a self-standing cognitive trait. Moreover, the specific mechanism for social learning developed by humans are themselves subject to cultural evolution (Heyes 2012b; 2016) and therefore hardly stands as a logical precondition to culture.

[1.3] And yet, here we are, communicating with formal languages about complex, abstract things through mass-produced computers engineered at one end of the world and manufactured at the other. This constitutes an awe-inspiring accomplishment, which no other known form of life (regardless of their intelligence or sociality) has ever approached. So what makes us distinctly human at the end? Based on comparative cognition, the most widely accepted answer is the robust human propensity to develop both the ability and willingness for pursuing socially shared goals, i.e. *shared intentionality* (Tomasello et al. 2005; Tomasello and Moll 2010). Critically, shared intentionality does not constitute a Big Special adaptation explanation of cumulative cultural evolution. It is not a a cognitive trait in the classical sense of the term, *i.e.* an ability afforded by our individual mind. It is a pattern of activity that we find ourselves spending our life doing. Shared intentionality does heavily recruit classical “higher” cognition functions such as mentalizing (thinking about what other agents want) and metacognition (thinking about how our own thinking fits the task at hand). However, it can’t be explained away by the simple genetic selection of those function at an individual level. Indeed, nothing about “higher” cognition entails shared intentionality, and shared intentionality does not strictly necessitates it. The ancestors of humans could in principle develop primitive forms of shared intentionality without a drastic evolutionary change in their cognitive architecture, just in virtue of leveraging their existent social skills to their new savanna environment. For reasons which will become clear later in this section, we think of shared intentionality as an ecological practice existing prior to the evolution of modern human cognition, and then enabled (or even propelled) the evolution of “higher” cognition by dramatically augmenting the developmental opportunities and selection pressure to develop it.

[1.4] But how could shared intentionality explain the emergence of full-fledged cumulative cultural evolution? By the account of Tomasello et al. (2005), shared intentionality is characterized by three increasingly demanding modes of synchronization between agents. First, agents need to mutually engage, in the sense that they interact and respond to each other’s actions. Second, agents need to conceive of a shared goal, and to monitor each other’s actions and perceptions to ensure that they share that goal indeed. Third and last, agents need to conceive of a shared plan to reach that goal, as well as understand both their own and each other’s role within this plan. Given this operational definition, it is straightforward how shared intentionality affords by construction an efficient mechanism for the transmission of complex cultural information. Indeed, let us assume that two agents share the intention to get one to learn a skill mastered by the other. The skilled agent will demonstrate the skill, while the unskilled agent will monitor their actions to understand the underlying plan. Then, the unskilled agent will in turn demonstrate the skill, while the skilled agent will monitor their actions to make sure they understood the underlying plan. This pattern of behavior, generally known as "teaching" or "instructed learning", provides a robust, reliable mechanism for the transmission of complex cultural information due to the teacher's active monitoring of the student's skill. As far as we know, this ecological practice is specific to humans, and therefore provides a candidate explanation for the human specificity of cumulative cultural evolution (Tomasello 2016). Critically, the practice of teaching (and more generally shared intentionality) is a social activity involving the coordinated participation of several agents, rather than a self-standing cognitive trait that may be explained through natural selection.

[1.5] However, it is critical to emphasize that once we have opened the Pandora box of cultural cognition (through shared intentionality or otherwise), nothing about the human mind can be fully decoupled from culture. Indeed, collaboration and cultural transmission around material productions afford the evolution of an open-ended arrays of technologies (Stout 2021), such as fire, spears, or clothes, which drastically improve the range of survivability for humans. It also affords the development of cultural knowledge, and institutions such as collaborative care for the young, with the same effect. However, it also necessitates ecological traits facilitating life-long learning at heavy metabolic and life-history cost, such as long REM sleep, big brains, delicate stomachs, dangerous gestation and long infantile periods, which makes humans extremely reliant on collaboration and cultural knowledge for their basic survival. As noted by Henrich (2017, chap. 3), the misfortune of lost European explorers / colonizers provided us with a simple natural experiment showing what happens when modern humans are left to their devices to navigate even mild environments without the appropriate cultural knowledge: they die, quickly and inexorably. Therefore, the continued existence of any society depends on its ability to maintain a sociocultural niche affording the development of locally adequate survival skills, and a mature population capable to develop, maintain, then teach those skills. Since all humans need this niche to develop to a mature state anyway, there is no reason why critical information for the development of typical mature human cognition should necessarily be embedded in the genotype, rather than the sociocultural environment. This aphorism only evokes a broader argument about the role of development and niche construction in human cognition, as found *e.g.* in Sterelny (2007; 2011). For our purposes, it is sufficient to point out that any given cognitive adaption (even those that are universal or critical to survival) can in principle be heavily embedded in the social and cultural niche – and that, in fact, they often do.

[1.6] Let us illustrate that statement with the example of a basic cognitive function, working memory. Working memory is a central element of "higher cognition", or more precisely *executive functions*, a somewhat loose collection of cognitive function enabling the complex self-regulation of behavior toward specific goals (Cristofori, Cohen-Zimerman, and Grafman 2019; Diamond 2013). Working memory plays a central role in the orchestration of executive functions, as it provides an information buffer for other executive functions like inhibitory control, cognitive flexibility, planning and reasoning (Coolidge and Wynn 2005). The capacity of working memory therefore constitutes a hard limit for the execution of complex tasks. In modern humans, working memory is significantly extended by the *phonological loop* (Baddeley 2003), or in a layman's term the process of mentally repeating to oneself contextually relevant information. The role of the phonological loop in enabling language acquisition has been studied in detail. However, this process affords a reciprocal relation: the development of complex language, both at an individual and collective scale, would considerably extend the capacity of working memory. The cultural evolution of more efficient language, or more language-intensive societies affording a more afferent use of the phonological loop, could therefore play a heavy role in explaining the evolution of a enhanced executive cognition. In particular, it would explain the sudden onset of behavioral modernity ~40k years ago (with Coolidge and Wynn (2005) finding no clear signs of enhanced working memory before 14k years ago), much after the evolution of anatomically modern humans. In any case, we have ample evidence of speech-based augmentations of working memory in contemporary humans, eg through distributed remembering (Sutton et al. 2010) or through the prevalent practice of remembering history or pathes through song (Chatwin 2012; Norris and Harney 2014). Therefore, we can't evade the fact that the very function of working memory is now embedded in language and cultural practices - regardless of what is precisely the evolutionary relation between the two.

[1.7] With these considerations in mind, let us turn back to the basic cognitive attitudes underlying shared intentionality. By the third criterion described in 1.4, shared intentionality entails that the agent understand and enact a specific role in the shared plan. This entails that the agent abstract their engagement with the world from their own embodied experience, and understand it in terms of a social expectation within a coordinated activity. In other words, they must *stand outside from* their direct sensorimotor context through an act of imagination, and consider instead their own activity within a broader social context. By reference to de Haan (2017), we will refer to this cognitive attitude as the *existential stance*. This term evokes the philosophical or religious discourse regarding the meaning of human life, but here it reflect the more modest meaning of its Latin etymology : *ex sistere*, to stand outside from. We do not argue that the existential stance is, *per se*, exclusive to humans, or that it is sufficient to explain the enculturation of human cognition. However, we do claim that the simple practice of reflecting on one's situation in a given material or social context does a critical part of the work, and does so in a way that is not fully appreciated by cognitivist approaches to the mind. First, we claim that the existential stance enables the integration of social norms and cultural skills by the agent through role-taking and similar processes. Second, we claim that it affords the development of abstract concepts embedded in the human social niche, and therefore enables the cosntruction and integration of complex social organization. Much more work will be necessary to properly set the context of this claim, and its significance in understanding the processes underlying the evolution of human culture. Beforehand, let us try to articulate at a narrative level how the existential stance affords novel possibilities for human cognition, as compared to non-reflective modes of social regulation.

[1.8] Human reliance on an adequate sociocultural niche comes with the necessity to regulate behavior so as to deter aggression and favor cooperation and mutual care. Participative activities which enforce a fine-tuned synchronization of embodied experience, such as dancing and music, afford the regulation of emotions as well as a strong sense of social bonding (Shilton et al. 2020). In particular, music afford the development of social bonding far beyond the scale afforded by grooming, sex or any mechanism for social bonding we share with other primates (Savage et al. 2021b; 2021a). However, purely embodied modes of social synchronization do not afford a robust transfer of coordination skill toward other, more complex tasks, or explicit sociocultural norms. The existential stance provides a mechanism for overcoming this limitation through the explicit inference of social expectations, and the monitoring of one's actions to enact those expectations. In the same way intentional affords the robust transmission of skills due to active monitoring by a skilled agent, mutual monitoring between socially motivated agents afford the robust synchronization of social norms within a social group. The existential stance therefore affords a sufficient grounding mechanism for the cultural evolution of social norms through domain general learning, as proposed by the cultural evolutionary psychology account of Heyes (2023). Those norms would then be enacted through reflexive self-monitoring, motivated either by social conformity or by norm internalization. To be clear, this statement is not meant as a distinction between two distinct causal pathways toward norm enaction. We will discuss in part 2 the specific mechanisms we propose for normative cognition, and how they dissolve the definition of specific motivations for norm enaction (and even the notion of "norms" as external entities which the agent can choose to enact or not). However, it highlights the existence of a deep relationship between the existential stance and the regulation of human cognition.

[1.9] The regulation of cognition, understood as the definition of goals together with the self-monitoring of cognitive activity toward those goals, are generally defined in cognitive science as the *executive functions*. Diamond (2013) defines three "core" executive functions, namely inhibitory control (ie the very exercise of self-regulation), working memory (as discussed above, an enabling factor for complex self-regulation), and cognitive flexibility (which, by the author's definition, can be assimilated to reflexivity recruited in the context of complex self-regulation). Without any doubt, the exercise of the existential stance is embedded in those core executive functions, in the sense is contingent on (and constrained by) the agent's executive control. Somewhat less evidently, executive control appears to be reciprocally embedded in the existential stance, and in human social regulation more generally. This perspective is most obvious for "higher-order" executive function. For example, the work of Sperber and Mercier highlights the social embedding of reasoning (Sperber and Mercier 2012; Mercier and Sperber 2011) - see also (Williams (2023) for a related approach in media studies. In their view, reasoning is evolutionarily grounded in the need to evaluate normatively other's motivations and defend one's motivations in the eyes of other humans, with problem solving ability emerging as a byproduct. This notion is notably supported by the prevalence of motivated reasoning (which is actively harmful to the epistemic role of reasoning), as well as the empirical finding that reasoning is more systematically and skillfully mobilized in an argumentative context (Mercier and Sperber 2011, pt. 2). We may additionally note that reasoning is embedded in the cultural practice of language, which makes it by construction a socially embedded phenomenon. But given that the same is true about working memory, and that working memory constitute a central buffer for executive cognition, the same holds for any other executive function.

[1.10] We should here stop to consider the larger role of language in self-regulation, in relation to our broader purpose of demonstrating the cognitive underpinnings and consequences of the existential stance. As noted by Clark (2006), language provide a material symbolism to communicate one's train of thought. By construction, the materiality given to thought provides a straightforward mechanism to objectivize it, to make thought itself the topic of further examination. Let us assume I want to trail down a path near my home in order to gather mushrooms. I will anticipate in no particular order the visual impression of the trail, the way my gait will adjust to the terrain to move swiftly and silently, the visual signs by which I find my way to my usual spot, the end goal of picking and bringing back the mushroom, and I will enact the behavior associated to that anticipation. However, the only thing that constrains me to actually follow with my plan over contingent environmental solicitations is the uninterrupted flow of my anticipations. Would any distraction break this stream of thoughts, I'd instantly forget my task and chase whichever new goal I am now imagining. Let us now that I repeat to myself, perhaps out loud or perhaps in my head, "I will trail down the path near my home to gather mushrooms". Now my plan stands outside me as a material object, a stream of sounds or of abstract symbols against which I can compare my behavior, and correct my course of actions were I to drift off path. Moreover, were I to decide to change plans, I could simply decide to direct my attention elsewhere by repeating my new plan to myself. Language, and more specifically self-directed speech, provides a straightforward way to externalize thoughts into symbolic objects with which to plan and reason, and against which we can constrain our action and attention [1].

[1.11] This discussion does not aim to speculate on the relation between language and executive control written large. It aims to highlight that unlike inner thoughts, linguistic expression can be directly examined and communicated to other. By using language, I may explicitly negotiate the nature of my goals with other agents, and then monitor my own activity against those goals. In others words, the words and sentence by which I constrain my actions are themselves up for public discussion – explicitly if I were to utter them out loud, and implicitly through the direct influence of other's public speech. Although language enables a lossless inter-individual transmission of specific units of meaning, which may reach arbitrary levels of complexity through composition, the integration of social and self-regulation is not in itself specific to language. It is rather specific to the existential stance. The practice of self-regulating by "standing outside" my own embodied perspective makes my own course of action an abstract object which I can mentally manipulate. As we discussed above, this practice is routinely mobilized in the interpersonal negotiation of shared plans and of each

1 This last sentence constitutes a straightforward reformulation of Clark (2006)'s argument. Although its motivations and ontology diverge wildly, it would be unfair not to note how it resonates with the work of Archer (2003) on the relation between structure and agency in human societies.

actor's role within those plans which underlies shared intentionality. By construction, the way I think of myself reflexively is therefore the topic of social negotiations, either directly or through the role others expect me to take. We could imagine a cognitive decoupling or dissonance between private existential reflection, and the public role I adopt through shared intentionality. However, a systematic decoupling between private and public life would be cognitively costly, as it would entail to maintain several parallel complex regulation modes, as well as the ability to contextually pick and switch between them. As we will see in the following sections, it may not even be possible due to the nature of the cognitive processes underwriting the perception of material and social context.

[1.12] We have hereby discussed at a conceptual level the evolution of cultural evolution, and more particularly of the complex ways in which human self-regulation has become embedded in social activity. We have accentuated the role of language in executive functions, as this illustrates most directly the fact that processes which we would intuit as properly cognitive are in fact embedded in cultural and social activity. We could also have outlined the role of material niche construction in canalizing human activity and extending our cognitive and organizational capacities. Arguably, language (as an instance of material symbolism) constitutes a prototypical instance of this process. The core argument we want to drive, however, is the key role played by the *existential stance* in the development of encultured cognition. The fact we can think of ourselves in reflexive terms (whether linguistic or not), and abstract our own experience to consider how we appear to others and how we participate in social processes changes entirely the possibilities for social coordination. Indeed, it opens the door for explicit patterns of social regulation, as we come to anticipate the expectations associated to a given role and situation beyond the reach of our direct experience. To be clear, those patterns do not only include "social norms" in the classical sense prescriptive rules of behavior, but also more intimate aspects of cognition such as self-directed language, emotional regulation or technological practices. However, this account remains until now entirely narrative, and is therefore a far cry from a positive, scientifically useful model of human cognition. To propose an actual framework for understanding encultured cognition and the possibilities it affords, we would need a specific formal model of how human cognition works and how it maintains and integrate "culture". We will turn shortly to providing exactly such a framework, which will explain the emergence of *social* *constraints* based on the account of the existential stance afforded by the Active Inference framework.

2) The Active Inference of cultural landscapes

[2.1] We will hereby expose the Active Inference Framework (AIF), a radical model of human cognition based on a physics-first approach, and its many implications for the cultural evolution of social regulation. Our central claim is that Active Inference (as a formal and conceptual framework), coupled with the cognitive process we defined as the existential stance, affords the emergence of rich modes of regulation embedded in the sociocultural and material niche shared by existential agents. We call this phenomenon *embedded normativity*, following Guénin--Carlut and Albarracin (2024). The concept of embedded normativity relies heavily on the non-separability of the cognitive processes underlying perception and action under Active Inference, both being driven by a single "self-evidencing" process where the agent acts so as to maximize the evidence for a world model which mathematically corresponds to its internal dynamics. Under this model, there is no meaningful separation between "descriptive" processes which underlies perception and "prescriptive" processes which underlie action. Both epistemic and pragmatic considerations shape at the same time the activity of the agent, from the percepts they are likely to form to the actions they are going to enact – including the epistemic action, such as eye saccades, they will undertake to gather information. Due to this non-duality, the regime of social regulation entailed by embedded normativity essentially trivializes the difference between "norms" in the classical sense of prescriptive rules, and non-prescriptive modes of regulation such as linguistic conventions or technological know-how. It also dissolves the principled separation between the agent's internal cognition and the structure of its world, in such a way that the structural properties of the shared material and sociocultural landscape (*i.e.* the specifics patterns of embedded normativity) becomes a proper explanation for individual behavior.

[2.2] A good way to introduce the AIF is to point out how uncannily efficient cognition is at adaptively constraining behavior. Indeed, to be able to navigate their world cognitive systems must be able to infer the causal structure of their environment from raw sensations. This process, which was termed "perceptual inference" by the pioneer of psychophysics Hermann von Helmholtz, has nothing trivial. Indeed, we only have access to a limited level of information (*e.g.* for vision, 2 2d array of pixel-like information corresponding to retina cells activity) which do not naturally correspond to the relevant causal structure in the world. To directly infer top-down representation of our surroundings would take the inversion of a "forward model" of how states of the world would expectedly shape our sensations, a task that is notoriously intractable even for simple models with few parameters. Not only does this process necessitates a cyclopean amount of computational power, but it typically constitutes a wildly unstable operation where even modest noise in the data causes extreme variations of the inferred model. That issue is further compounded by intrinsic noise in biological computation, by multi-modal integration (*e.g.* integration of visual and auditory signal), by the cost of adjusting the model in real time, and by the physiological limits on information flow which can go up the brain hierarchy to be integrated in high-level models. Simply speaking, not enough information leaves the optic nerves to ground a top-down representation of our surroundings, if there was enough it wouldn't be sufficiently accurate to constrain it to be anywhere near the actual structure of the world, and if it was accurate enough it would be much too slow to enable real-time adjustment of the model. To this day, we simply don't know of a biologically plausible or even physically possible way to solve that problem.

[2.3] A solution to this conundrum is, counter-intuitively, to rely more heavily on the background knowledge of the agent to circumvent the necessity to inverse the forward model – or even to compute it accurately in the first place. Let us assume that instead of trying to brute force the structure of its environment based on sensory information, an agent tried to predict its sensations by assuming a given forward model of the world and adjusting it in real time. This computational architecture has many advantages over the one outlined above. First, the intrinsic computational cost of inversion disappears entirely, allowing for much faster (real-time) adaptiveness of the perceptive model with biologically plausible metabolic costs and information flows. Second, because the model need not be inverted and can be adjusted in real time, its accuracy stops to matter immensely. This means that biological noise cease to be a problem for the robustness of perceptual inference, and that aggressive simplification can be used to further diminish the computational and informational cost of perception. Third, the integration of information through time, space, and sensory modes into robust and meaningful perceptual experience, which compounds all of the difficulties of perceptual inference, becomes a somewhat trivial matter of maintaining coherence in background knowledge. If your model supposes that a specific object may cause different modes of sensations, or maintain itself through space and time, then your perception of the object shares those properties by construction. Fourth, efficient generalization of conceptual structure is built in the basic structure of predictive perceptual inference. If I observe that a cat mews, then I'll come to expect that things that look like the cat mews – and this without necessity for any background knowledge of the notion of "cats".

[2.4] The different theories which postulate that human cognition follows this broad predictive architecture are collectively known as Predictive Processing, which is exposed and discussed in its earlier formulation in Clark (2013; 2015). Some of the best empirical arguments for Predictive Processing are very easy to experience, even without a laboratory or expert knowledge. Let us ask the reader to keep their eyes on the page while shifting their attention towards their peripheral vision, to try and perceive their surrounding environment. The reader should be able to reconstruct a pretty complete picture of their surroundings, including the capacity to "experience" the perception specific objects, but remain unable to retrieve any of the specific details that should be necessary to reconstruct the nature of those objects. For example, the reader may see a chair in the corner of their eyes, but fail to retrieve even basic information like its general shape or the presence of armrests. It is rather easy to play on your expectation to bring you to see things that are not there – you could for example keep "seeing" the chair at the corner of your eye if someone moved it while you were not paying attention. Actually, the very fact your visual field feels complete relies on this effect. Indeed, a significant portion of your retina happens to be obscured by your optic nerve, but you still experience vision in the corresponding portion of your visual field do to this effect. This is because our peripheral vision lacks the precision to see anything but vague movement and shapes, but our visual experience is "constructed" through the predictions our visual neurons come to integrate. Computationally speaking, this boils down to the fact our retinal cells' activity seem to encode deviation from expected signal, rather than raw sensory information (Hosoya, Baccus, and Meister 2005).

[2.5] Applying the same basic insight to nested scales of neuronal organization leads to the Hierarchical Predictive Coding (HPC) framework, which constitutes the most widespread version of Predictive Processing. Indeed, it explains many otherwise puzzling empirical facts about human cognition and neurocomputational architecture. For the reasons exposed above, the very fact we are capable of reconstruct in real time a robust and informative perceptual experience from sensory stimuli is best explained by

HPC. The same features that enable HPC to easily account for the seamless aspect of perceptual inference extend to the predictive inference of interoceptive states (*e.g.* gut sensations), therefore explaining the intimate integration of body and emotion in "higher" cognition (Seth 2013; Corcoran and Hohwy 2017). Perhaps most importantly, our ability to generalize linguistic or concept structure from very little information constitutes one of the most important puzzle in the history of cognitive science, and it is essentially explained away by HPC (Tenenbaum et al. 2011; Tenenbaum, Griffiths, and Kemp 2006). HPC is additionally well-tuned to account for the variety of human phenomenology and ethology. For example, the phenomenology and the patterns of behavior associated the diagnosis of depression map well onto habitual stress response under HPC (Badcock et al. 2017), while those associated with Autism Spectrum Disorder and Schizophrenic Spectrum Disorder map respectively onto increased strength of ascending and descending signal respectively, as compared to neurotypical population (Tarasi et al. 2022). More generally, the misattunement of predictive mechanism easily grounds the development of inaccurate, self-confirmed belief or percepts, matching the ethology of psychosis (Sterzer et al. 2018) and dementia (Kocagoncu et al. 2021). Finally, while it remains insufficient for any definitive conclusion, neurophysiological evidence is overall coherent with HPC (Walsh et al. 2020). In particular, the central role of descending prediction signals in HPC explains the prevalence of descending connections in most brain areas, which fits poorly with any model based on direct (non-predictive) inference.

[2.6] The Active Inference Framework (AIF), which this article calls onto, constitutes a variant of HPC which posits that cognition is organized around a specific design principle: the minimization of Variational Free Energy (VFE) (Friston, Kilner, and Harrison 2006; Friston 2010; Friston, Daunizeau, and Kiebel 2009). Simply speaking, the underlying narrative is that cognitive agents strive not only to predict their environment, but also to maintain the conditions in which it is predictible – by actively looking for information through actions such as ocular saccades and exploration, but also by acting onto the environment to reduce variance at its source. Starting in the late 2000s, a series of publication by the British neuroscientist Karl Friston have suggested that the brain functions by minimizing VFE (Friston and Kiebel 2009; Friston 2010; Friston, Kilner, and Harrison 2006), a common metric of model fitness used in statistical inference which constitutes a upper limit of surprise (formally speaking, the probability of the observation made by an agent). Friston has then moved to proposing the minimization of VFE as the process underlying self-organization in living systems (Friston 2012; 2013) then (more recently) articulating a general theory of self-organization as active inference (Friston 2019). This line of research have led to the result known as the Free Energy Principle (FEP), which states that any system coupled to (but distinct from) its environment minimizes VFE (Da Costa et al. 2021). Minimization of VFE entails that the internal dynamics of an agent maximize the Bayesian coherence[2] between action-oriented, minimal, predictive expectations (formally, Bayesian priors) about its environment and its actual sensorimotor experience. The FEP constitutes today the central claim to legitimacy of the AIF, a relation which is best articulated in Parr, Pezzulo, and Friston (2022).

[2.7] Under Active Inference, both perception and action become embedded in the same imperative for maximizing evidence for the generative model of the world entertained by the agent (Friston, FitzGerald, Rigoli, Schwartenbeck, and Pezzulo 2016; Friston, Parr, and Vries 2017). This process is sometimes referred to as "self-evidencing", as the generative model of the agent is formally dual to the dynamics which underlie its existence (Hohwy 2016). Therefore, the AIF has succeeded at constructing a formal account of cognition which is grounded in the dynamics of the cognizing life form (M. D. Kirchhoff 2018; M. Kirchhoff et al. 2018). In doing so, it has constructed a middle ground solution to the long standing debate around the role of "representation" in cognitive science. This debate stems from a contradiction at the very foundation of cognitive science: the formal model of the human mind it proposes was based on logical operations made on symbolic representations on the world (as a computer would), but its model of brain activity was based on dynamical systems and did not afford grounding such symbols. This led to the development of two conflicting proposition: one being that we should abstract the "hardware" of the brain to focus on the "software" of the mind and its supposed language of thought (J. A. Fodor 1975; J. Fodor and Pylyshyn 1988) (see also Piccinini (2010) on the motivations and meaning of this metaphor), the others that we should forget about symbolic representations and focus on the dynamics relating body, mind and environment (Brooks 1991; F. J. Varela, Thompson, and Rosch 2016). Active Inference therefore dissolves the need for *bona fide* symbolic representation by grounding the semantics of cognition in the very dynamics

2 "Bayesian" refers here to Bayes' rule, the mathematical formula underlying optimal statistical inference about a state A given the observation of a related state B, a prior belief about A p(A), and a forward model p(B|A): $p(A|B)=\frac{p(A)*p(B|A)}{p(B)}$

that constitute it (M. J. Ramstead, Friston, and Hipólito 2020; M. J. Ramstead, Kirchhoff, and Friston 2020; Constant, Clark, and Friston 2021).

[2.8] Finally, regardless of its theoretical and conceptual virtue, Active Inference has its merits as a self-standing computational theory of human cognition. Indeed, extending predictive processing to decision, action, and planning solves the same kind of computability problem that it does for perception (Hipólito et al. 2021), with the same kind of ground-breaking implications. As noted in Clark (2020), a central feature of this account is the dissolution of "agency-as-we-know-it", in conflict with our folk intuition about agency [3]. Under Active Inference, motivations and intentions are described as prior belief over the states an agent expects to observe. As action (like perception) is driven by the minimization of VFE (*i.e.* model evidence), agents will act so as to enact those priors (unlike it's less costly to revise them). This account realizes the core function of "motivations" or "intentions": there are cognitive states which orient action toward specific outcomes. However, it displace the locus of agency from the internal mental states of the agents toward the broader attunement between the agent's dynamics and its environment's. As discussed in Bruineberg and Rietveld (2014), the prototypical Active Inference agent experiences the world as a *landscape of affordances*, where the term "affordance" refers to opportunity for action as directly perceived by a cognitive agent. The predictive account of agency entails that those affordances are perceived as immediately normative, as one anticipate themselves acting on them and assess compatibility with their anticipated sensori-motor flow. Assuming you stand over a cliff, you would likely not consider at all the possibility of willingly walking to your death. If you did, you would be struck by vertigo, and (assuming you want to live) feel the need to distance yourself from that possibility by walking away from the cliff. In other words, you'd perceive the very anticipation of falling as negative, rather than consider this course of action neutrally and then deciding not to act on it.

[2.9] To describe this phenomenon, we can talk of *embedded normativity* (Guénin--Carlut and Albarracin 2024). It refers to the fact that agents under Active Inference do not experience the norms and values guiding their actions as the expression of an intrinsic

3 Please note that this comment is directed toward contemporary, Western-dominated academia. For all relevant purposes, the concept of the mind as standing above and regulating matter was inherited from Greek cosmology, and reified further reified in Christianity and the Enlightenment (Lent 2017). Chinese philosophy, for example, is on the contrary well-tuned toward a minimalist model of agency, as the nature of *action without action* or more idiomatically *effortless action* is one of the core axis of its Axial philosophy – as discussed in Slingerland (2014).

will, but as the property of landscape of affordances they experience. Norms and value may be embedded in "internal" percepts, such as interoception or self-directed talk, but they remain embedded in the basic fabric of predictive experience. This phenomenon becomes especially meaningful when we consider the implications for existential agents such as humans. As a reminder, such agents are characterized by their capability of representing themselves reflexively, especially in the context of interpersonal coordination around collective plans (*i.e.* shared intentionality). Here, we turn back to highlight that adopting the existential stance in the context of interpersonal coordination entails the construction of (and is driven by) shared *regimes of attention* (M. Ramstead, Veissière, and Kirmayer 2016; Veissière et al. 2020). Indeed, shared motivations, shared intentions, and *a fortiori* enaction of shared plan entails that agents have a common model of what is relevant to their common pursuit, and where their attention should be directed. But in doing so, they constrain the landscape of affordances they experience to specific percepts and actions that are coherent with the shared regime of attention. Hence, they produce *cultural affordances* - immediate perception of opportunities of action that are embedded in locally prevalent cultural patterns. This logic may as well apply to overt skills (such as crafting or hunting) as to covert skills (such as emotional regulation), and even to skills that reflexively apply to how skilled intentionality is enacted (such as language or metacognitive social learning rules). In other words, the construction of cultural affordances under existential Active Inference is sufficient to ground the radical enculturation of human cognition we have described in part 1.

[2.10] A question we need to address is how robust exactly is the transmission of normative cultural patterns under Active Inference. Indeed, the way a predictive agent can understand and engage with its world is inherently reconstructive, and this applies as well to social norms or cultural affordances (Guénin--Carlut 2022a). Even if a group of agent may reflexively associate the same cultural affordances to the same context, the landscape of affordances which they concretely experience may vary drastically. Most obviously, one's experience of shared hierarchical norms vary immensely with one's position in the hierarchy, with straightforward effect on the stress levels who are consistently cast in a subservient role (Rivers and Josephs 2010). More subtly, variation in one's interoceptive experience, or in the material affordances associated in strength and height, or in the skill level relative to such and such tasks may lead to different patterns in existential reflection, and therefore to reconstructive drift in the social norms integrated. The robustness in the transmission of social norms and cultural affordances would be greatly augmented by the existence of even one anchoring system, which may carry specific percepts without loss between agents. It just so happens that we have one: language. As discussed in Clark (2006), language serves as an anchoring system that can carry thoughts both within and between agents. We may highlight here how Active Inference offers a straightforward way to embed language in all dimensions of cognition, therefore supporting the tentative account of executive functions made in previous section. More specifically, it affords what Karl Friston calls *sophisticated inference* (Friston et al. 2021), ie recursive forms of planning based on beliefs about beliefs. Language indeed affords to externalize, communicate and manipulate explicit beliefs, be it about the state of the world, one's or other peoples beliefs or intentions, or normative expectation in a given context.

[2.11] However, it is more urgent for our argument to stress that externalizing thoughts into words «is to create an object available to ourselves and to others[, which] we can have thoughts about.» (Clark 2006). In other words, to enact language is to shape one's social niche by communicating specific expectations, and therefore constraining one's and other's behavior (Tison and Poirier 2021; Tison 2022). This provides a powerful mechanism for the robust transmissions of cultural affordances, especially for existential agents and in the context of their role-taking activities. Indeed, language may embed arbitrarily complex *social scripts*, understood here as complex patterns of behavior associated to the performance of a given social role in a given cultural context. As analyzed in Albarracin et al. (2020), scripts constitute complex instances of cultural affordances, which can be accounted within Active Inference in much the same way given the ability to represent such scripts. We should highlight that scripts are especially effective at constraining behavior when they co-evolve with rigid system of social identity, such as gender or ethnicity. Indeed, expecting "an agent such as oneself" (as formulated in Veissière et al. (2020)) to perform a specific script entails (given that we do identify with this agent) that we expect ourselves to perform that script. This prompts agents to dissipate prediction error, typically (assuming both the identity and associated norms are strongly embedded in the social niche) by performing the script Guénin--Carlut (2022b). Therefore, the intimate integration of normativity and self-identity through scripts enables the robust transmission and production of social norms and cultural affordances through engagement with a given cultural landscape. This provides us with a proof of principle that there are at least a structural isometry between the landscapes of cultural affordances experienced by different agent, in the sense that there exist universal constraints over expected behavior (applying both to oneself and others).

[2.12] We have hereby shown that the existential stance is sufficient to ground the development of a radically encultured cognition, assuming that the Active Inference model of cognition is essentially correct. This is because the abstraction of an agent "such as oneself" drives the agents to predict given patterns of attention or behavior in both oneself and others. Once reified and stabilized, those patterns become cultural affordances, and they collectively constitute a cultural landscape of affordances which participative to produce humans' direct experience of their world. Critically, a cognitive agent is considered under Active Inference to be driven by the very structure of the cognitive landscape they experience. Therefore, cultural normativity is embedded within the cultural landscape which humans collectively enact, rather than being a matter of rational choice or decision over a perceived space of possibilities. The coherence (or more modestly the inter-intelligibility) of the landscapes experienced by different humans in a same cultural ecology is ensured through an array of synchronization mechanism. The most notable is language, which produces an "objective" reference reality that all agents in the local ecology (or a significant majority thereof) may in principle access and manipulate. However, embodied synchronization through the experience of a common material niche or through dancing and ritual activity also offer important leverage for the coherence of the shared cultural landscape. Although we consider it a proof of principle for our central thesis, this account should raise an array of question regarding the dynamics and phenomenology of cultural landscapes. Most notably, it remains unclear how human activity is shaped by cultural landscapes, and how those landscapes evolve through time. Although a full account of the underlying formalism is beyond the scope of this paper, we will turn to the physics of Active Inference to ground the notion of a *social constraint* and explain how this specific concept informs the present account.

3) A gauge theory of social constraints?

[3.1] We have exposed above how (given the model of cognition entailed by Active Inference) existential agents such as humans integrate and enact landscapes of cultural affordances, giving rise to the dissolution of agency into a regime of embedded normativity. What we aim to do now is to motivate the existence of a specific construct, the *social constraint*, as the basic brick of embedded normativity as well as a proper *explanans* for (but also a description of) aggregated patterns of social behavior. Depending on the context, the term of constraint can either refer to *formal constraints* (mathematical properties that a system or collection of systems must verify by law or construction) or *causal constraints* (physical entities which causally shape given processes without being affected by them at the relevant time scales). We draw on this dual meaning to accentuate the fact that, in the context of our discussion, they have a common ontology. A social constraint *is* the social regularity that can be observed in a given cultural context, but it also *is* the force that draws locally encultured agent to respect this regularity. The duality between those two aspects of social constraints derives directly from the nature of agency under (existential) Active Inference: assuming a given agent believes a given regularity characterizes their social niche, they're going to expect themselves to follow it and therefore be drawn to do so. We will expose the physical nature of the underlying forces, and discuss the conditions to the coherence of a shared cultural landscape. Overall, we will formally motivate (although not fully explicit) a physics of social construction and lay out its core properties. These formal considerations will then allow us to reframe the conceptual considerations introduced so far into an original account of social agency and enculturation.

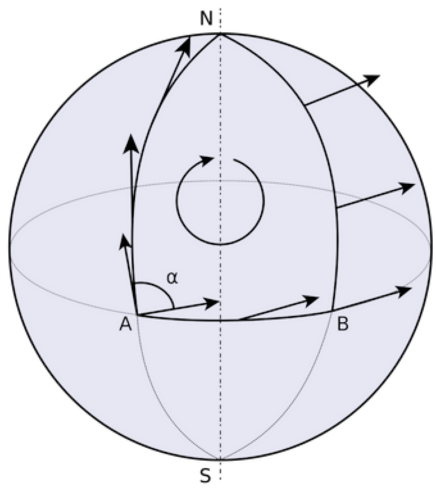
[3.2] As discussed above, the term of "constraint" can refer to two different constructs, recruited in the context of two different explanatory strategies. The concept of *formal constraint* refers to mathematical properties that a system (or a collection of systems) must verify given *a priori* physical principles, known boundary conditions of the system, or conditions of its construction. In this context, constraints can serve as prior knowledge in the search for mechanism explaining given phenomena, or they can serve to directly explain existing relations between function and structure (Green and Jones 2016). For example, formal constraints can constitute scaling laws relating mass and metabolic rates in animals (Green 2015) or similar relations between the spatial and energy scale of oscillation in materials, but it can also constitute the use of specific mathematical technique to limit the search space in the resolution of an optimization problem. By contrast, *causal constraints* are actual physical structures which causally shape the target of explanation so as to limit its range of behavior (Ross 2023). For example, the barrel of a cannon may constrain the degree of liberty of the gaz liberated in detonation to force the movement of the cannonball, or the structure of road network may produce specific patterns in the commercial and demographic exchange between cities (Fulminante et al. 2017). Importantly, the two notions of constraints are very different in their nature and scope, but they're not mutually exclusive. Indeed, causal constraints are by construction structures which are not affected by the target of explanation, and which do not change at the relevant timescale (Montévil and Mossio 2015). Therefore, formally accounting for a causal constraint in the description of a system means (at least in the most straightforward cases) writing down a formal constraint. The difference between the two notion relies here on what aspect of a same system is being considered: its formal properties, or its causal implementation.

[3.3] When we talk of a "social constraint", we may therefore talk of the formal properties verified in a given social context, or we may talk of the structural factors that bring about those patterns. This dual meaning would normally be a source of confusion, and necessitate a clarification of what specific definition we give to the term. We believe on the contrary that it reveals the dual nature of constraints as functional patterns and as structural causes. In other words, social constraints can be characterized at the same time as a properties of invariance in aggregated pattern of social behavior, and as the causal forces which realize those pattern. The meaning of that statement is less than trivial, and exposing it in adequate details will take some diving into mathematical physics. More specifically, we can explain the nature of social constraint by analogy to *gauge theories*, a type of field theories where a specific type of physical regularities (*gauge symmetries*) manifest from the observer's frame of reference as the emergence of corrective forces enforcing those regularities (*gauge forces*). We may understand gauge symmetries as fundamental facts of nature which we are (most of the time) not equipped to observe directly, and gauge forces as the way those symmetries appear to us. The case of social constraints is similar in the extent that the conservation of specific expectations over social behavior will appear to an external observer as an *ex nihilo* causal force over behavior. However, it differs deeply in the sense that the symmetries being conserved are constructed by the very activity of the system subjected to the forces, rather than constituting a preexisting physical fact. With this limitation in mind, let us attempt to expose as clearly as possible the physical meaning and the motivation of the "gauge forces" account of social constraints.

[3.4] Imagine two persons play ball standing on a rotating platform. The conservation of momentum by the ball means that it will travel in a straight line, but from the perspective of the rotating players it will seem that it is deviated to the side as it travels. In other words, from the rotating frame of reference, we can only account for the conservation of momentum by adding a corrective force which is taken to produce the deviation. This force is called the *Coriolis force*, and it constitutes a very simple illustration of how gauge theories work. A gauge force fundamentally emerges from a misalignment between the formulation of the problem and the frame of reference in which it is observed. Were the observer to place themselves in any inertial frame of reference (*i.e.* in any point traveling in space at constant speed without rotation), they would observe the ball traveling at constant speed and direction, and therefore respect conservation of momentum. This is because in any inertial frame of reference, the basic dynamics of the problem (the speed and direction of travel of the ball) is conserved. It is only because the rotation of the platform causes the observer to travel along a path where the dynamics of the problem are not conserved that they observe a curling in the trajectory of the ball. To be clear, this does not imply that the Coriolis force is simply an useful fiction. It has very real, measurable effects if and when a system is forced to travel along non-inertial frames of reference. For example, because the atmosphere and water bodies on the planet Earth happen to be traveling on the surface of a rotating sphere, they are subject to the Coriolis force with observable consequences on global meteorological patterns. The same logic apply for other, less intuitive fundamental physical regularities.

[3.5] In general, the most basic property of physical theories is the collection of *symmetries* verified by the system. Symmetries represent a family of transformations over the problem's formulation which do not change its dynamics (or more precisely, do not affect the Lagrangian operator the dynamics must minimize in accordance to the principle of least action). For example, the ball throwing problem described above can be described in any inertial frame of reference without reformulation of the dynamics, which means that it is symmetric with regard to Galilean transformations. Collections of symmetries typically have the structure of a *group*, which means that they include the identity transformation (which is trivially true), that their element can be composed and inverted, and that they also include the product of such operations. Such symmetries can be discrete, like for example the symmetries induced by the reversal of space (P: $\vec{r} \to -\vec{r}$) or time (T: $t \to -t$) dimensions. They can also be continuous, like the Galilean symmetry of our ball throwing example (T: $(\vec{r};t) \to (R \circ \vec{r} + \vec{r'} + \vec{v}\,t; t+s)$) corresponding to the change between referential frames of reference. Each symmetry can be traced to observable properties of the system, and in particular any continuous symmetry corresponds by Noether's theorem to the conservation of a given quantity through time (Noether 1971). For example, symmetry with regard to translation in space entails the conservation of momentum. Indeed, if an object is accelerated through its trajectory in space, then its dynamics become dependent from where it is observed in space, as specific regions will be associated to quicker movement [4]. In other words, translation symmetry breaks down. What is critical to understand here is that, while we have derived earlier the existence of Coriolis force from conservation of linear momentum, the relevant physical quantity (as far as gauge theory is concerned) is the underlying translation symmetry.

[3.6] The power of gauge theory, however, comes from the case of *local symmetries*, where the gauge transformation apply independently to any point in space. The space of gauge parameters can then be framed as a bundle of *fibers* extending independently from each point of the *base space* in which the theory is framed. The fibers are related to other nearby fibers by a *connection*, which specify how the gauge parameters change along with infinitesimal variation in the base space. We can illustrate the meaning of a connection with the simple case of parallel transport over a sphere (see figure beside). In that context, we have a vector that is tangent to the surface of the sphere, and that we move along said surface while conserving its orientation in the tangent space. As is clear on the figure, the idea of "conserving the orientation" of the vector only



Parallel transport of a vector around a closed loop (from A to N to B and back to A) on the sphere. From Wikimedia Commons.

4 To be clear, this explanation is only tangentially related to the canonical demonstration of Noether's theorem, which relies on the fundamental properties of the Lagrangian. See (Baez 2020) for details.

makes sense as we move it along an infinitesimal length, as the curvature of the sphere causes the orientation of the vector to change along a closed loop. The connection specifies the rule by which we need to change its orientation in the total space to account for the change in the tangent space as it moves along the surface. By analogy, and even though the choice of gauge parameter is free at any point in spacetime, the connection provides a geometry connecting gauge parameters associated to infinitesimally close points in the base space. Much like the vector in our example appears "accelerated" as it travels across a curved geometry, paths that travel along curved gauge geometry (*i.e.* paths that deviate from the connections) appear as accelerated by a *gauge force* in the base space. This mechanism generates all of the fundamental forces known to physics [5]. For example, when we talk of an object being accelerated by the electromagnetic force, we are really referring to travel along a path which is curved in the geometry induced by the U(1) symmetry of Schrödinger's equation (see *e.g.* Hooft (2008) for a in-depth treatment of gauge theories and their role in contemporary physics).

[3.7] For the purpose of the present discussion, we will not need a formal understanding of the mathematics underlying gauge theories - at least not beyond the intuition provided by the simple example above. The core intuition we will build on is that phenomenon we experience as fundamental physical forces *really are* curvature in the geometry induced by fundamental symmetries of physics (along with the rule of propagation they entail). Critically, the same logic applies to the physics underlying Active Inference. As we discussed in part 2, Active Inference is motivated by the Free Energy Principle, which states that dynamical systems coupled to but distinct from (in a specific statistical sense) their environment perform approximate Bayesian inference by maximizing coherence between expectations about its environment and its actual sensorimotor experience. We will now discuss the formal details underlying this statement. The core condition under which the FEP applies is the partitionability of a

5 Although gravity can be understood as a gauge force produced by the local geometry of spacetime, it is not typically described as a gauge force as defining a quantized theory of gravity remains an open problem. However, the quantization of gauge theories constitutes a much later development of physics, and its importance falls well beyond the scope of the present argument. We are using here the term of "gauge theory" in reference to theories that are constructed to be invariant with regard to a change of basis. This is coherent with original formulation of gauge theory, which (as recounted in O'Raifeartaigh and Straumann (2000)) focused specifically on the argument that vectors or fields can only be compared locally, and any physical theory should therefore be invariant with regard to a change of basis at any given point of spacetime.

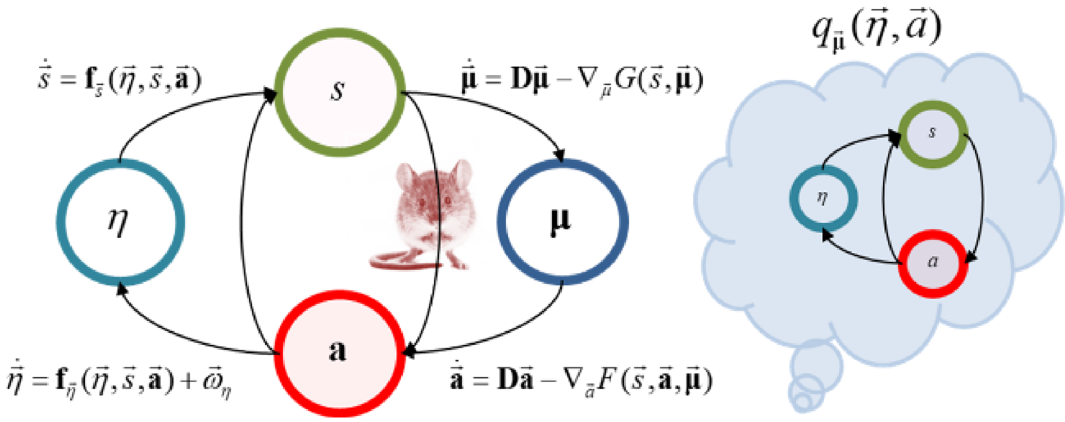
stochastic dynamic system into two subparts separated by a "Markov Blanket", *i.e.* a set of state which mediates all statistical dependencies between the other two subparts (Da Costa et al. 2021) [6]. When this conditions is verified, there exist an information geometry which associates each internal state of an arbitrary non-boundary subpart (taken as the agent) to a distribution over its complementary (taken as the environment) (Parr, Da Costa, and Friston 2020). Furthermore, the internal flow of the agent minimizes Variational Free Energy for this distribution. This is equivalent to say that it performs near-optimal Bayesian inference over the space of beliefs implicitly entailed by the information geometry. Critically, this statement can be reframed as a gauge theory entailed by the symmetry of the system's dynamics with regard to reparametrization of its statistical attractor (Sakthivadivel 2022; M. J. Ramstead et al. 2022).

[3.8] In this Bayesian gauge theory, the drive toward Bayesian coherence (which we described qualitatively in part 2, and constitutes the formal motivation for Active Inference) emerges as a gauge force driving the system toward higher probability states. To be precise, the dynamics of the system can be recast in two orthogonal components: a vertical flow toward a *synchronization manifold*, *i.e.* a subregion of the space corresponding to the most likely states given the state of the boundary, corresponding to the gauge force ; and an horizontal flow within this synchronization manifold, corresponding to the enaction of Bayes-optimal inference and planning by the agent. Let us take some time to recollect what this statement means, concretely. The horizontal flow, by construction, is *conservative* – it conserves the measure (the "size") of the subspace the system has access to. This means it can encode the kind of quite complex dynamics we can expect of living and cognitive systems. More specifically, it can underwrite the cyclic behavior (sleeping, foraging, hibernation...) underlying homeostasis, as well as the dynamic bifurcations necessitated by decision making (active as well as perceptive). But conservative dynamics are also, in a fundamental sense, antagonistic to adaptive control. By definition, adaptive control aims to *reduce* the measure of the subspace a system has access to by constraining the controlled variables to expected outcomes, while conservative dynamics *conserve* this

6 Be (A, B, E) a partition of the stochastic dynamical system S. B is a Markov Blanket if and only if $p(A|B, E) = p(A|B)$ and $p(E|A, B) = p(E|B)$. This is equivalent to say that the distribution of both subsystems are independent of each other, when conditioned on the distribution of the Markov Blanket. Critically, this condition is different from the absence of direct dynamical coupling between A and E.

measure. This is where the vertical flow comes into play. In a Bayesian gauge theory, it works to constrain the conservative flow in the specific region of the synchronization manifold, where it corresponds to the specific ecological cycles and decision tasks enacted by the agent. In colloquial terms, we could say the Bayesian gauge force *canalizes the chaos* to the specific dimensions where it can drive the information processing and adaptive engagement with the world underwriting an agent's cognition.

[3.9] While the theory may explain *why* cognitive systems display apparent attraction toward adaptive control, we still have to demonstrate *how* it works at the level of a given agent. The core result we need to build from is that, while any statistically individuated system induces the existence of a Bayesian gauge theory, the type of complex behavior we associate to cognition is specific to particles which Markov Blankets can be partitioned between sensory states (which are not affected by internal states, and do not affect external states) and active states (which are not affected by external states, and do not affect internal states) (Friston et al. 2023). In substance, this is because active states then becomes embedded in the space over which the information geometry encodes expectations. Those expectations can then encode complex predictive "beliefs" over one's own actions, which produce decision making, motor control, and planning. But then, the only outside constraint on the internal dynamics of an agent - which underwrite the evolution of their beliefs - is the stream of sensations they're exposed to. Given this, the only mechanism that may correct deviations from optimal beliefs (*i.e.*



Architecture of strange particles. Only sensory states affect the dynamics of internal states, and only active states affect those of external states. From Friston et al. (2023)

the most likely internal states, as dictated by the information geometry) is to adaptively select which sensations will be used to update one's beliefs. Incidentally, this confirms the earlier treatment of Active Inference as a neuronal gauge theory, which explicitly designated attention as a gauge force emerging from the symmetries of sensory entropy (Sengupta et al. 2016; Sengupta and Friston 2016). Indeed, attention is modeled under Active Inference as the precision of specific predictions, which computationally translates into the weight given to prediction error in the updating of beliefs. Together with (and augmented by) salience, *i.e.* the intrinsic motivation to perform a given epistemic action, attention provides the computational filter enabling agents to select which sensations will be experienced and allowed to influence their own internal flow.

[3.10] For this very reason, producing specific constraints over the flow of attention may offer a powerful ground to fundamentally transform the space of cognitive possibilities accessible to a specific agent. In particular, let us recall the cognitive attitude we defined in part 1 as the existential stance, *i.e.* the reflexive attendance of one's own activity in third person perspective. While it may seem (and is probably) a peripheral, culturally embedded component of human cognition, it also enables agents to evaluate their own actions against the same expectations they apply to other agents engaged in similar tasks. In other words, it affords the *abstraction* of specific sensori-motor schemes from their private, embodied instantiation. An existential agent expects themselves to act as one ought to do in their given context, and their sensorimotor flow is hence driven to realize this expectation. Conversely, the expectation of how one ought to do in a given context is driven to be coherent with the embodied experience of the agent in the same context. Due to the very reflexive attendance to one's actions, Bayesian gauge force enforces the coherence between abstracted expectations applying to anyone and the private, embodied expectations that drive behavior. Incidentally, what we are describing here is only a reformulation of the cognitive psychology of enculturation, as it was described in part 1. Reflexive self-monitoring affords the negotiation of task-specific roles between co-operating agents, and therefore the emergence of "complex" social learning through the integration of complex sensori-motor schemes embedded in those roles. In other words, it constructs a common ground enabling agents to share their private knowledge about a task by comparison against a public cultural standard for how the task should be realized. Hence, assuming Bayesian gauge theory is correct, the existential stance is sufficient to explain the evolutionary constructions of human cognition through cumulative culture.

[3.11] Hence, reframing Active Inference as a gauge theory enables a powerful formal motivation of the conceptual account presented in part 2, which holds that the bulk of cognition is (under Active Inference) implemented through systematic constraints over the flow of attention, and enculturation emerges from the coupling between the attention of different agents. The core contribution of Bayesian gauge theory is to demonstrate the duality between the formal symmetries characteristic of the beliefs it enacts, and the causal power of attention in maintaining the coherence of those beliefs. In other words, social constraints exist simultaneously as a (gauge) symmetries over the system's statistics and as (gauge) force over its dynamics. In a fundamental sense, they *are* at the same time *formal* and *causal* constraints. This statement is, as exposed earlier in this section, a core feature of the present account. To believe *is* to observe (and engage with) the world in a constrained manner, and to observe (and engage with) the world in a constrained manner *is* to believe. By extension, the participation of several agents in an ecological niche characterized by common constraints over attention is sufficient to define a common ground of belief-like attitudes, *i.e.* a cultural landscape of affordances. However, the specific manner in which Bayesian gauge theory accounts for this duality fundamentally limits its ability to reproduce key components of cultural phenomena. Indeed, physical gauge theories stem from *a priori* formal constraints, entailed by fundamental gauge symmetries. By contrast, the cognitive expectations that Bayesian gauge theory aims to explain are constructed through the contingent activity of the very same agents who enact them, as is manifest in human cultural evolution. Hence, there is no clear way by which it seems to explain the open-ended evolution of social constraints, as opposed to the simple evolution of a well-defined system in a predefined space.

[3.12] We have hereby exposed the formulation of Active Inference as a Bayesian gauge theory, *i.e.* a fundamental physical theory relating the symmetries of a system to the causal force driving its activity. Here, it demonstrates the duality between the structure of one's expectations and the systematic way in which one's attention flow so as to maintain the coherence of those expectations. It would seem that framing encultured cognition in this formalism would straightforwardly provide a principled account of enculturation as a physical process. It does provide a specific physical model (the only one to our knowledge) of how encultured engagement with the world *work*, namely by driving the way an agent projects salience and meaning onto specific stimuli so as to realize social constraints. But at the same time, the formal grounding of Bayesian gauge theory does not afford to explain how encultured engagement with the world *develops*. Indeed, the dynamical system-theoretic framing of Bayesian gauge theory presupposes that the system follows a predefined, fixed dynamical flow. On the contrary, agent understand the cultural landscapes they evolve in through a reconstructive process, in which their characteristic features may evolve by undergoing loss or creation of information. In other words, if we can describe the enaction of social constraints as a gauge force over attention, we still have to explain how those constraints are contextually integrated and (re)constructed by cognitive agents. We propose to close this explanatory gap by exploring further the process of constraint integration, which we understand as a (re)constructive process through which agents actively brings about (and reshape) the very same system of social constraints they evolve in. Specifically, we will explore the idea that cognitive contextuality in the perception of cultural landscapes offers subtle but critical nuances over pure dynamical systems theoretic accounts of encultured cognition, which enables to construct a coherent theory of open-ended cultural evolution.

4) Becoming human : contextuality, social semiotics, and participatory realism

[4.1] We have so far articulated a formal theory of the enaction of social constraints which is grounded in Bayesian gauge theory, a formal theory of cognitive behavior. Bayesian gauge theory formalizes the idea that agents engage with their environment through the continuous anticipation and attunement of sensori-motor flow. Simply speaking, the core idea is that social constraints correspond to the geometry of a shared landscape of material, social, and cultural affordances which the agent is brought to engage with through their development. However, as noted above, this theory fails to reproduce important features of cultural phenomena. By construction, a given cultural landscape (unlike a presumably fundamental physical field) exists only through its perception and enaction by encultured agents. Encultured agents do not have a direct access to the "objective" structure of this landscape, as this structure simply does not exist. They can only reconstruct a "subjective" version of a shared system of constraints, as inferred from the public behavior of its neighbors. This process is intrinsically reconstructive, as the agent needs to fill the ambiguities in the definition of this landscape (be they intrinsic or relative to its perspective). But the constraints as experienced by the agent actually shape their behavior, which is in turn perceived by other agents and used to infer the structure of the cultural landscape. Therefore, the specific way in which a given agent "subjectively" understands a given system of constraints is a constructive force in its "objective" structure. In other words, cultural landscapes are by construction the subject of a continuous and active reconstruction by their participants, which matches the empirical observation of open-ended evolution of human culture. This does not match the general structure of the theory of cultural engagement entailed by Bayesian gauge theory, which fully specify the possible states and trajectories of the system *a priori* .

[4.2] Let us rearticulate this claim in more formal details. By definition, a dynamical system-theoretic object is defined by the triplet (T, X, Phi), where X is the *embedding space* containing the possible states of the system ; T is the time dimension ; and Phi is the flow function Phi: (X, T) → X. The flow function describes the movement of the system in time given its initial position, which entails the coherence conditions Phi(x, 0) = x ; and Phi(t_1 + t_2 , x) = Phi (t_1, x) · Phi(t_2, x) for all x, t_1, t_2. By construction, a dynamical system can only flow in and from the specific embedding space it is defined in. The definition of novel dimensions in which the system can evolve is simply not an element of this theory. Of course, proponents of Bayesian gauge theory could argue that the information-theoretic or geometric aspects of this theory enables it to transcend the limitation of dynamical systems theory. While the *definition* of novel dimensions is not described within the embedding space of the theory, their *recognition* by an agent or their *emergence* through active engagement with their world could (in principle) be described through the construction of an information geometry of the problem. However, the information geometry in question is defined from the marginalization p(·|b) of the attracting distribution p(·), itself being defined by Phi. By construction, this means that all the formal constructions underlying the system's belief, cognition, and individuation are fixed *a priori*, by the very act of defining the embedding space and the flow of the system (Guénin--Carlut 2022c; 2023). In consequence, the theory is structurally unable to describe the perpetual redefinition of the formal or causal constraints characterizing a given system through its own activity, and by extension the (re)constructive aspect of enculturation.

[4.3] Therefore, describing how cultural landscapes evolve through time and enable the construction of genuinely novel possibilities will necessitate a redefinition of our approach which goes (subtly but critically) beyond the scope of Bayesian gauge theory *per se*. In fact, it will necessitate an important re-articulation of what we expect mathematical representation to accomplish. Indeed, the classical way to address a problem in physics is to define the set of states it can possibly realize (*i.e.* define its phase space) and articulate a law which *entails* its evolution at the relevant scales (Longo, Montévil, and Kauffman 2012). Here, this approach won't help much, as the question being addressed is precisely how the embedding space and evolution law *come to be* (Guénin--Carlut 2022c; 2023). As noted *e.g.* in Kauffman and Roli (2023), this requires a theory framed outside the classical grounding theory for mathematics, set theory. This is because, by construction, sets constitute well-defined spaces which are fully specified by the collection of states they include. A set can only include a given state or not, and a function defined over a set can only give one result to one object. In other words, set theory assumes the individuation *a priori* of mathematical objects and their properties, and provides a questionable ground to represent objects whose properties are fundamentally observer-dependent, such as social constraints. This is not to say that mathematical representation is impossible, as other grounding theories of mathematics (*e.g.* topos theory) may provide a better account of observer-dependent objects or properties. However, our focus remains hereby on deriving useful theoretical insight for social and cognitive sciences. We will therefore sideline fundamental coherence and strictly formal considerations, and stick to describe what kind of physical process is the (re)construction of social constraints.

[4.4] Let us recollect what properties of sociocultural landscapes of affordances / social constraints makes them difficult to represent mathematically. First, the properties of a sociocultural landscape of affordance do not constitute observer-independant physical properties, but they are defined by their very perception by a given (ensemble of) participating agents which is embedded in the material, social and cultural context of its engagement with the aforementioned landscape. We call this property *contextuality*. Second, the space in which those landscape are defined is itself a contingent product of system dynamics, which therefore escapes embedding in a predefined space. We call this condition *open-ended evolution*. Critically, those two properties are not independent traits of sociocultural landscapes of cultural affordances / social constraints which could be described separately. They constitute two different manifestations of the fundamentally reconstructive nature of human culture. Therefore, a successful theory of enculturation would need to account for the fundamental relation between contextuality and open-ended evolution. This is precisely where the duality between the formal and causal account of social constraints comes into play. As formal symmetries underlying an agent's engagement with the world, social constraints are to be understood as the geometry of the agent's dynamics in an objective world. But understood as causal constraints over one's attentional regime, social constraints are precisely what brings about the space the agent can perceive and meaningful engage with. This gives a pivotal role to the specific framework of constraints: when holding the dynamics of an agent's engagement with the world as constant, we may describe the construction of social constraints as the result of "active inference" ; but when holding the constraints embedded in an agent's engagement with the world as constant, they are the basis of the space in which active inference occurs. This enables, at least in principle, the description of *open-ended evolution* of the actual dynamics of the system as the outcome of *contextuality* in the states an agent perceives.

[4.5] Again, to provide an adequate formalism to account for this process is beyond the scope of this specific paper. Our goal here is to point to the physical process underlying encultured cognition to argue that the intuition is in essence correct. To do so, the core claim we need to support is that social constraints over encultured agents' attention can bring about novel, contextually meaningful observables ; and that this process shapes the space of possibilities which is effectively experienced and navigated by those same agents. As we discussed in [1.8-10], language constitutes a very direct illustration of that logic. At the barest level, words constitutes material tokens for specific units of meaning, whose value is entirely embedded in the education of human attention to look for those specific units of meaning in those specific material tokens. Straightforwardly, their communicative value relies on *contextuality*. But at the same time, it grounds the genuinely novel possibility to externalize those specific concepts and organize them in a syntactic manner. Therefore, their development entails the *open-ended evolution* of sociocultural landscapes of affordances. When anticipating a given course of action, the agent may also anticipate having to explain themselves to other linguistic agents they communicate with, or may simply talk to themselves about it. This entails the coupled attendance to their immediate sensori-motor flow and to the words or sentences which describe them. Therefore, the very afference of linguistic actions induces novel constraints over the agent's dynamics of engagement with the world, defined by the internal structure of language. In this sense, the development of language does not simply add new dimensions to the prior dynamics of the agent, but feed back into the very geometry of its non-linguistic engagement with the world. The same logic applies more generally to the various semiotic systems that a community of encultured agents may use to communicate with each other.

[4.6] Indeed, a community of agent is encultured in that they generally agree on the specific *meaning* of specific *signs* embedded in social activity. This is precisely this shared agreement which mediates the existence of social and cultural affordances, as encultured agents can then infer what other encultured agents meant by specific acts. Those signs can be full-fledged symbolic tokens of meaning, like words, but they can be more subtle communicative acts relative to rhythm, poise, or tone. For example, the teaching of technical skills relies on the successful communication of the syntax of a given task (which may be symbolically represented) but also on the reconstruction of the specific sensori-motor flow which correlates with the successful completion of the subtasks (which may not be symbolically represented). Barring telepathy and similar parapsychological processes, the only way the teacher can communicate this second information is by relying on specific signs that are understood by the student. For example, the teacher may intentionally overemphasize certain elements of their movements to help guide the student's attention toward the specific parameters that are critical to the completion of the task. In this case, while the ultimate sensori-motor schemes which underlie the technical skill are ultimately private, they are developed through (and understood in relation to) the given system of sign that mediated the communication between the teacher and the student. In other words, the ability to infer the private intention of public acts can be explained from the direct engagement with a world affording communicative acts, regardless of the preexistence of a *bona fide* theory of mind. What matters, for the specific purpose of our argument, is the basic attitude of abstracting specific actions from one's direct sensorimotor context and to considering their value as *communicative acts*. In other words, it is what we have called the existential stance.

[4.7] With this in mind, we may turn back to the mosaic account of the development of encultured cognition we proposed in section 1. Once educated to understand their own actions as communicative acts, existential agents may use those systems as a way to scaffold their cognition toward complex, socially integrated modes of regulation. The abstraction of specific concepts into symbolic or quasi-symbolic communicative acts enables an intuitive comprehension of how the existential agent's acts are to be understood from the outside, of how they play out in the context of collectively negotiated norms and intentions. I may communicate with others what tasks I intend to take on, whether I expect them to help, and how. Hence the public negotiation of private intentions toward a given collaborative activity is constitutive of shared intentionality, which we have introduced as the main ground of human cultural cognition according to *e.g.* Tomasello et al. (2005). Our discussion of Active Inference, and of the importance of contextuality in cultural cognition, enable us to provide a more specific model of how shared intentionality works and interacts with human cognition. In particular, we claim that shared intentionality (and therefore enculturation) is mediated by the integration of human cognition in the context of social semiotics, through the coupled anticipation of material engagement with conventional tokens of meaning. This provides a solid ground for the mosaic account of human cognition as we have presented it in section 1. Indeed, its main theme is the notion that social regulation may somehow be turned inwards and drive the integration (and open-ended evolution) of complex modes of cognition. This seems quite straightforward, if we agree as a premise that engagement with social semiotics is dual to the integration of the adequate regimes of attention to perceive those semiotics.

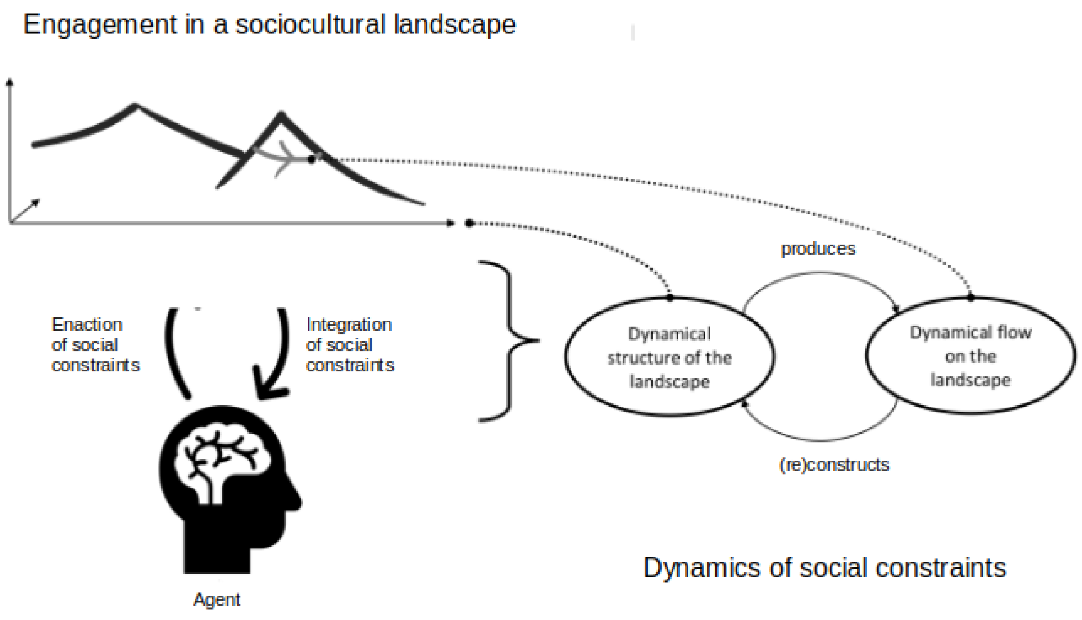
[4.8] Indeed, an encultured agent can understand the private intentions of their own actions in relation to the public communicative acts they instantiate. Why do I go to the river? To fetch water, probably. Abstracting those specific concepts from one's immediate sensori-motor flow entails an extreme contraction of the cognitive complexity of articulating abstract goals and intentions which use them as a primitive. I may simply bring myself to "go to the river in order to fetch water for Grandma to cook with" by recalling this sentence, instead of having to hold in mind the entire collective sensorimotor sequence in which the simple act of walking toward the river is integrated. Here, we may turn back the hypothesis articulated in section 1 that executive function is related to the integration of language in cognition. In our view, cognitive engagement with a field of conventional signs carrying communicative meaning is a plausible grounding for the three core executive functions of inhibitory control, working memory, and cognitive flexibility (Diamond 2013). Simply speaking, conceiving one's activity as oriented toward an abstracted goal articulated in socially relevant signs enables the agent to monitor their own activity toward that goal (hence inhibitory control), to retain abstracted tokens of information relevant to that goal as the context of their activity (hence working memory), and to use the compositional aspect of social semiotics to situate their activity in complex, contextual sequences of action (hence cognitive flexibility). Embedding complex sequences of action in a synthetic language is a sufficient mechanism to canalize the agent's attention toward the specific cues that are relevant to those actions, and therefore bring themselves to enact those plans. However, we cannot emphasize too much that, by construction, embedding social semiotics in an agent's cognitive engagement with the world makes the regulation of one's attention a collaborative exercise.

[4.9] Indeed, the world of signs that an agent engages with are their own production as much as it is this of their neighbor, parents, priests or urban planners. Guénin–Carlut and Albarracin (2024) has described this phenomenon through the dual concepts of *internalization* and *externalization*. By engaging adaptively with their material and social context, encultured context learn to *internalize* specific way of attending to the world which enable them to perceive the contextually relevant social signs or cultural affordances. But this internalization corresponds to the *externalization* of the adequate signs / affordances, in that they come to perceive those contextual objects as objective features of the world. In the extent that a cultural community agrees on the basic make up of social signs / cultural affordances (*i.e.* that they have *internalized* similar constraints over their attentional flow), encultured engagement with a material and social context can then be treated as a shared information channel, where agents may *write* as well as *read* in symbolic states. One may expect encultured agents to understand precisely which of those signs they generated themselves, and abstract the process as a simple perceptive input. But as the reader recalls, the Active Inference model of cognition hardly affords such a separation between perceptive domains. According to our discussion in section 2, it is precisely the blending between perception and action entailed by their coupled anticipation which explains the uncanny ability of Active Inference agents to understand and engage adaptively with their world. If a given act is more coherent with the social context and identity in which an encultured agent evolves, if it is more likely to bring about the states of the world the agent expects, it is more likely to become afferent and to be enacted.

[4.10] This simple fact has many consequences regarding the way that specific norms or cultural expectation can be embedded in one's material and social environment. By *externalizing* a given language or semiotic system and then using it to *write* on the world, an agent may drive their comrades to learn to *read* their message and then *internalize* the logic of their language. Imagine a group of Roman citizens is so impressed by the military power of their empire that they come to compare it to the Sun ; they may then imagine the cult of Sol Invictus, together with a religious symbolism and ritual activity which embeds their beliefs and transfers it to converts. The same symbolism may then be imprinted in a cyclopean stone construction, which architecture is meant to reflect the beliefs of the cult through direct figuration or indirect symbolism. Now, passer-by can be direct witness to the power of the cult as well as its logic, and may become more inclined and more capable to engage co-operatively with it. However important material niche construction is in human social organization, one other domain of social semiotics has a privileged relation with action and decision-taking : the domain of self-identity. Indeed, an agent may come to understand themselves through a given culturally afferent role such as a *teacher*, a *student*, a *woman*, a *man*, a *parent*, a *child*, or a *magistrate*. Then, the likelihood of their own actions become directly conditional to the abstracted nature of their social identity. Of course, the afference of such roles is conditional on a social and material context which affords it: a *parent* cannot parent without a *child*, a *magistrate* cannot rule over the dust. But pragmatically, those roles condition the expected response to specific social signs for the agent, and hence the nature of the cultural affordances they experience.

[4.11] This discussion enables us to ground the narrative account of enculturation articulated in section 1, while providing much needing detail about the nature of the underlying processes and its implications. If our argument is in essence correct, enculturation derives from the contextual nature of social semiotics and cultural affordances. Indeed, the semiotic nature of specific material signifiers does not lie in (although it may derive from) their intrinsic physical nature. It resides in the shared understanding of this signifier as a communicative act meant to convey a specific meaning, in the context of a community of encultured agent sharing (to some extent) a system of communication. Due to the connection between attention and phenomenology entailed by Active Inference, we may explain how socially and materially embedded constraints over the agent's attention leads to a shared experience of specific signs as intrinsically meaningful. In other words, we can describe enculturation as a process in which agents *internalize* specific regimes of attention so as to *externalize* specific signs as (perceived) objective states of the world the agent engages with. However, the reader may recall that the FEP assumes the identity between the perceptive concepts deployed by the agent and the actual external states of the world. In consequence, if we are to describe it in terms of the FEP, externalization does not only correspond to the redefinition of the dynamics of the system, but also of the very space in which the system evolves. In other words, we may reconstruct the *open-ended evolution* of encultured cognition from the *contextuality* of social semiotics. Of course, this argument relies critically on the notion that there exist social semiotics in the first place – *i.e.* that social activity may produce abstracted signs, whose relevant meaning lies in the perspective of the agents who perceive and enact them rather than their bare sensori-motor dynamics.

[4.12] While it may not be necessary for the development of semiotics in general, the existential stance entails the basic attentional disposition to treat given acts as communicative in nature, and infer an intended meaning abstracted from their bare sensori-motor dynamics. This amplifies in order of magnitude the ease with which agents may agree on abstracted social semiotics, and the complexity which those semiotics may reach. Therefore, the constructivist logic we have applied to language applies as well to less formalized domains of communication, and it does so in virtue of the role of the existential stance in human cognitive development. But in doing, it also amplifies the importance of purely contextual properties or objects in human ecology. The idea that object or properties in the world can in fact derive from their contextual perception by physical observer or cognitive agents is know as *participatory realism* (Fuchs 2017; Guénin--Carlut 2022b; Froese 2022; Guénin--Carlut 2023). The specific philosophical implications and formal premises of participatory realism are as of now mostly uncharted, and they are also out of the scope of the present article. However, we do need to outline that it entails the fundamental of *which observer* a given state is defined for. Specific cultural concepts may exist as an actual causal force in driving social activity, but only does so because specific agents have developed the prior attentional dispositions which bring about those concepts as meaningful elements of their world. Therefore, social constraints do not simply act as causal forces over an agent's attention. They also act as elements of an interpretative system which defines the field of social information – the very field in which social constraints are in turn defined. We will now turn to the question of whether we can study the dynamics of social constraints for their own sake, abstracting their material implementation in human activity, and what are the consequences of this perspective.



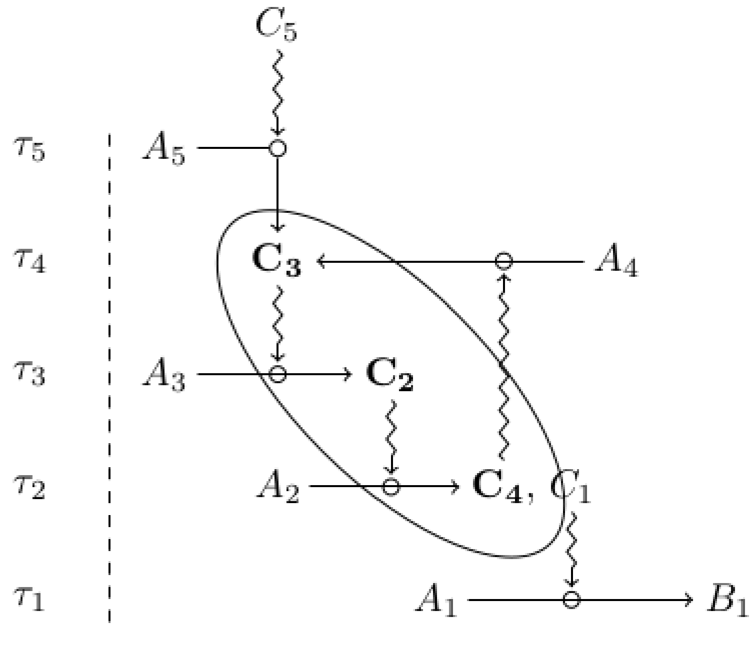
Reconstructive engagement with a cultural landscape – from Guénin--Carlut and Brenac (2023)

5) Constraint closure and the emergence of sociocultural lifeforms

[5.1] In the present article, we have proposed that "social constraints" embedded in human perception-action cycles constitute the core driver of enculturation. We have used the Active Inference model of the mind, and its formal counterpart Bayesian gauge theory, to demonstrate that the (encultured) patterns embedded in one's attentional flow (*i.e.* one's "beliefs", in the minimal sense proposed by Active Inference) must indeed constrain one's sensorimotor engagement with the world. We have furthermore argued that this model can explain the *enaction* of social constraints, but cannot explain their *evolution* or *integration*. Indeed, the static background of Dynamical System Theory allows for the definition of *a priori* formal constraints, but not for the adaptive and contingent development of causal constraints throughout a system's life. We have finally proposed a minimal mechanism for the open-ended evolution of sociocultural landscapes, in the form of a semiotic model of constraint integration. In this model, the development of constraints over attention is dual to the *externalization* of a system of signs through which the agent perceives and engages with the world. Collectively, these arguments demonstrate that social constraints *exist* in a non-trivial form, and that modeling a given sociocultural landscape as a system of constraints (continuously reconstructed through engagement with social semiotics) is *coherent* with the goal of accounting for the open-ended evolution of human encultured cognition and social organization. We must now explain why the existence and coherence of social constraints opens the door for a novel picture of human social organization as a *bona fide* lifeform, capable of self-construction. Drawing from Wittgenstein (1953)'s vocabulary, the rules (implicit or explicit) for social activity happen to constitute a *form of life* – understood here as an organization that meets the criterion for biological organization. We will now expose this thesis, and explore its further consequences for the relation between *structure* and *information* in the multi-scale dynamics of human societies.

[5.2] The formalization of what constitutes *life* is a longstanding and important question in the modern scientific tradition. Schrodinger (1944) correctly anticipated it to be grounded in the reproduction of symbolic information (which later was proved through the discovery of DNA), while Prigogine and Nicolis (1971) situated in the context of the thermodynamics of dissipative structures. A synthesis notion of life as a self-sustaining network of processes emerged in the work of Stuart Kauffman, especially Kauffman (1969; 1993). This notion was soon reappropriated by foundational work in the enactivist school of cognitive science, most notably through the notion of *autopoiesis*. Autopoietic organization was initially defined as "a network of productions of components which (i) participate recursively in the same network of productions of components which produced these components, and (ii) realize the network of productions as a unity in the space in which the components exist" (F. G. Varela, Maturana, and Uribe 1974). The study of cognition was recontextualized through its role in the "realization of the living" (Maturana and Varela 1991), *i.e.* how cognitive activity participates in the conservation of the autopoietic organization that constitutes the cognitive agent. The literature surrounding autopoiesis was conducive to an important change in the study of life and cognition, but suffered from two major flaws that prevented its broader acceptation as a self-standing theory of life. First, the way in which it tried to explain away the teleology of life effectively prevented the description of adaptive behavior: an autopoietic organization either conserves its structure or it dissipates, it has no room for change or adaptation (Di Paolo 2005). Second, it characterized biological organization at the relatively abstract level of "operational" properties, which was effectively decoupled from the description of the concrete physical structure and processes that drive life (Ruiz-Mirazo and Moreno 2004; Ruiz-Mirazo, Peretó, and Moreno 2004).

[5.3] As a consequence, enactivist literature as a whole moved towards the formalization of an array of notion that were initially peripheral to autopoiesis, and that would integrate at a deeper level the teleology inherent to biological organization. In particular, the approach pioneered by Ruiz-Mirazo and Moreno (Ruiz-Mirazo, Peretó, and Moreno 2004; Ruiz-Mirazo and Moreno 2004; see also Moreno and Mossio 2015) proposes that "'a living being' is any *autonomous system* with *open-ended evolutionary capacities*" (Ruiz-Mirazo, Peretó, and Moreno 2004, 330) (emphasis added). The concept of basic autonomy is specified as self-construction, understood as the ability for a given system (*e.g.* an autocatalytic enzyme network) to generate its own structure). It differs from the earlier "top-down" autopoiesis approach by its explicit focus on the thermodynamics of biological processes (inspired notably by (Kauffman (2000)'s work-constraint cycle) and its proposal to “step forward and get wet, specifying what kind of constraints—and associated types of work—are needed to effectively achieve minimal self-constructing systems” (Ruiz-Mirazo and Moreno 2004, 241). In other words, the concept of basic autonomy as formalized by Ruiz-Mirazo and Moreno centers the definition of life around the mutual relationships between *concrete* physical constraints which collectively drive *concrete* physical processes into (re)constructing themselves, thus driving the individuation and construction of living systems. This concept was later refined through Montévil and Mossio (2015)‘s concept of biological organization as a *closure of constraints*. This approach is virtually identical to basic autonomy in its implication and broad philosophy, except that it explicitly focuses on the functional relationship between the specific constraints and processes constituting biological organization. It can therefore articulate an explicit, formal definition of biological organization as a system of constraints with *closure*, *i.e.* where “each constraint plays the role of both generative and dependent constraint” (Montévil and Mossio 2015, 186).



Biological organization as a closure of constraints (Montévil and Mossio 2015; Guénin--Carlut and Brenac 2023)

[5.4] In the context of the broader organization, constraints are not defined by their physical structure *per se*. They are defined in virtue of their role in shaping the relevant process so as to co-produce the organization which produces them. In other words, constraints are only defined as such through their function in and dependence on the broader organization: “we claim that constraints subject to closure constitute biological functions. Within this framework, performing a function is equivalent to exerting a constraining action on an underlying process or reaction. […] Closure is then what grounds functionality within biological systems: constraints do not exert functions when taken in isolation, but only insofar as they are subject to a closed organisation.” (Montévil and Mossio 2015, 186). This relational and functional understanding of biological organization brings back the very problem that earlier approaches to the definition of life had circumvented by avoiding teleological or functional language: the question of *how* functional organization relates to the underlying structure. However, it also builds the capacity for open-ended evolution into the very definition of biological organization. Indeed, the evolvability of novel traits require the accumulation of genetic or structural variation that does not impede the functional organization of the system (A. Wagner 2012; 2008; G. P. Wagner and Zhang 2011). Conceptualizing biological organization as functional constraints over physical processes equips (by definition) the organism with the ability to transform its own structure (*i.e.* accumulate variation) in the very process of maintaining its own organization (*i.e.* maintain function). To cite again Montévil and Mossio (2015, 190): “As biological systems undergo functional variations, their organisation maintains closure, albeit realised in different variants, because of the continuous acquisition of some functions, and the loss of others. […] Understood in this way, the invariance of closure may be said to be complementary to its functional variation”.

[5.5] Let us revisit that argument in more details, using the framework developed in the part 3 and 4 of the present article. A given constraint subject to closure just is *what is given form* by a given process in the architecture and *gives form* to another process. In other words, describing biological organization as a closure of constraints only specifies the constraints under consideration in the *information* they carry about in the context of the functional architecture of the broader organism (in the sense of Jablonka (2002)). Following Guénin--Carlut, White, and Sganzerla (2023), we may consider in information-theoretic terms a constraint subject to closure, as a message that the organism sends to itself. This intuition may be formalized further through Bayesian gauge theory. Indeed, it entails that to a given structure (formalized as a stochastic dynamical system partitioned by a Markov blanket) corresponds an information geometry, which parametrizes the belief (or information) about the world carried by the internal states of the structure. Assuming a given constraint is implemented in a subsystem C which is individuated from the broader organism by a Markov blanket, the process leading to the production of a constraint by the broader organism can be described as Active Inference by the organism onto the constraint. The role of the organism in giving form to the constraint can therefore be understood as the externalization of internal beliefs over expected states. Conversely, the role of the constraint in producing the organism can be considered as the internalization by the organism of beliefs embedded in the structure of the constraint, in the same way complex attitudes are embedded into the material and social structure of semiotic systems. As discussed in part 4, this process induces construction of a novel semiotics defining what parameters matter in the overall interaction, and therefore is a candidate driver of open-ended evolution [7].

[5.6] This argument should help expose the strength and weaknesses of Montévil and Mossio (2015)'s account of life. On one hand, their focus on the functional and relational aspect of biological architecture adds another, arguably excessive, layer of abstraction between biology and the underlying biophysics. This runs directly counter the objective of *grounding* biology in the underlying physics of its substrate, as is the explicit goal of earlier approaches (*e.g.* Ruiz-Mirazo and Moreno (2004)). On the other hand, it forces the theorist to adopt the viewpoint of *life itself*, as they must consider the substrate as something that must be acted on to maintain the organism but whose intrinsic structure is ultimately unknown. For the same reasons we have articulated in section 4, the fundamental ambiguity between the "language" of the organism and this of the material substrate it is embedded in provides provides room for the emergence of open-ended evolution. Material or developmental variations that are not legible at the level of the organism may become relevant due to a change in structure or context, and hence become integrated in its functional organization. In principle, this provides room for the production of the *adjacent possible* (Kauffman 2014), and therefore for tendency of life to escape legibility by entailing laws and embedding spaces (Longo, Montévil, and Kauffman 2012; Kauffman and Roli 2023; Guénin--Carlut 2022c). Perhaps most importantly, we should note that Montévil and Mossio (2015)'s relation definition of life is by construction abstracted from the specific substrate in which life is embedded. As such, it applies without loss of generality to an organization made of

7 Please note that the account of Montévil and Mossio (2015) only demonstrate that a closure of constraints is *compatible with* open-ended evolution. The claim that it *explains* it is a novel claim, which is relative to the specific framework presented here.

constraints in human cognition. If there exist social constraints on human behavior, and if some system of those constraints collectively display closure, then this system constitutes a form of life *stricto sensu*.

[5.7] We have demonstrated the existence of social constraints in part 3, where "social constraints" are understood (by analogy to gauge theories) as a force born from cultural expectations which drives the dynamics of attention (and therefore behavior) at a very fundamental level. Therefore, we only need to discuss the validity of the closure condition. Conveniently, its validity is made somewhat evident from contextual nature of encultured cognition. Indeed, for a given cultural ecology to subsist beyond one generation, it does not only need to be developed by a given population, but also transmitted to its descendants. This transmission is conditional on the integration by younger generation of the adequate constraints to perceive and engage adequately with the locally adequate social semiotics and cultural affordances. We cannot verify that the specific dynamics of each agent's attentional flow is identical to their elders, since the specific attentional flow developed by an agent is conditional on their specific situation in a given cultural ecology. However, it is sufficient to note that the vertical inheritance of any single social signs or cultural affordance requires that the adequate social constraints underwrite their externalization / internalization were indeed integrated by the new generation. In turn, this integration requires the existence of a broader material and social niche which canalize agents toward it. In other words, the coherence of a cultural context across generation requires the existence of a system of constraints, which collectively drive human attention so as to reconstruct each of its single components. Hence, the very coherence across generational change of a cultural context, together with its embedded social semiotics and social organization, is evidence for the closure of the system of constraints it underwrites. Therefore, even if we cannot necessarily identify every single constraint involved, any given cultural ecology corresponds at some level of description to a *bona fide* biological organization (again, in the specific sense of Montévil and Mossio (2015)).

[5.8] That conclusion should not be very surprising: human societies display the same kind of complex organization contingent on a metabolic flow of matter, energy and information, together with the ability to reproduce their organization. Most of the arguments that explicitly seek to separate cultural from biological processes seem to focus on the Darwinian aspect of individual-to-individual transmission (Mulder, McElreath, and Schroeder 2006; Claidière, Scott-Phillips, and Sperber 2014). While social learning is definitely non-Darwinian, due to the prevalence of horizontal transfer and the absence of well-individuated units of transmission, most of evolutionary process aren't either. "Soft" (*i.e.* non symbolic) systems of inheritance, like human social learning, are extremely prevalent in life overall (Jablonka and Lamb 2008; 2006). So is the horizontal transfer of information, although it may be somewhat atypical for animals (at the exception of occasional but important instances of viral transmission (Soucy, Huang, and Gogarten 2015)). Indeed, Darwinian inheritance and selection only constitutes a model of *some* instances of evolutionary change, rather than a fundamental requirement of evolutionary theory (Pigliucci and Muller 2010; Kevin N. Laland et al. 2015). In addition, if human social organization is the relevant unit of biological organization, information systems evolving in this context should be understood as mechanisms of regulation of the emerging collective metabolism, rather than full-fledged systems of evolutionary inheritance. Indeed, the well studied process of evolution transition generally associates the evolution of cooperation between evolutionary units (in this context, human individuals) to the development of information systems able to regulate those relations at the scale of the emerging organization (Jablonka 1994; Szathmáry and Smith 1995; Jablonka and Lamb 2006; Szathmáry 2015). In this context, systems of "hot" communication which are the basis for most human cultural transmission (including full-fledged symbolic language, which basic units remain subject to subjective interpretation and evolves through usage) are closer analogues to early RNA networks than to the very stable and heritable DNA.

[5.9] In consequence, the central question we need to address is not the existence of social lifeforms *per se*, but their relevance to the study of human societies and cognition. The main, most straightforward insight from this perspective is that social activity participates (by construction) in the self-creation of sociocultural lifeforms (Guénin--Carlut 2022e; Guénin--Carlut, White, and Sganzerla 2023a). For example, Guénin—Carlut (2022) proposes to consider the progressive transition in human history toward States societies in integrated urban systems as the result of the externalization by administrative systems of their implicit expectations about social structure. Most characteristically, the maintenance of a specialized administration and of urban systems critically necessitates the extraction of grain, the only type of food that affords to sustain a dense population, from a captive rural population (Scott 2017). Therefore, the extension of cities and States always comes with the production of a lifestyle reliant on grain production, of agriculture landscapes, and of a social structure which drastically limits the ability of commoners to travel or otherwise escape taxation. This historical pattern is understood as the externalization by administrative systems of specific expectations about social structure into *social constraints* embedded in the material niche and the social activity of their population. Although those expectations are understood to flow from the system of constraints constituting the sociocultural lifeform under consideration, this account specifically proposes a top-down approach centered on the agency of centralizing institutions and on the limitation of the information systems they use. This is coherent with Stewart (2020)'s "manager theory" of evolutionary transition, but also with Jablonka and Lamb (2006)'s emphasis on the role of information systems in evolutionary transition. The process of constraint production is then brought back to the strategies that States deploy to understand the world they exist in, and to create a world that is *legible* and *manipulable* through the limited window of the information systems they rely on.

[5.10] Guénin--Carlut, White, and Sganzerla (2023) attempts to discuss the possibility of contemporary scientist to infer the functional architecture of past social lifeforms based on the archaeological landscapes we have access to today. The fact that social lifeforms are embedded in the patterns of behavior and attention of encultured humans entails that their full architecture can be reconstructed by human individuals through engagement with the local sociocultural landscape. However, the nature of any given constraint is contextual to the activity entailed by the entire system of other constraints, some of which are embedded in patterns of social interaction we cannot possibly access. Therefore, there is no reason that we could accurately reconstruct any single constraint participating to a past sociocultural lifeform without having direct access to its online activity. However, Guénin--Carlut, White, and Sganzerla (2023) note that the constraints that constitute a lifeform must somehow be communicated to the agents. Hence, the dynamics by which social constraints are *embedded* in a cultural ecosystem and then *integrated* by agents therefore constitutes a fundamental choke point on the ability of a social organization to (re)produce its own architecture. They further specify that the production of social constraints may either rely on either of two processes: *structural embedding*, the production of actual material structure that enforce or enable the expected behavior (like roads or city walls); or *semantic embedding*, the production of signs that elicit the expected behavior in the acculturated population (like a road sign or a throne). Because of the role of structural embedding, the role of the material and social niche in transmitting those constraints may be reconstructed through careful modeling, especially regarding patterns relevant to visual attention and spatial navigation. This provides us with an avenue for investigation that has the potential to greatly augment our knowledge about the role of embodied phenomenology in social organization, especially in relation to the cognitive science of religion and material engagement.

[5.11] Conversely, this line of research exposes important areas of obscurity in existing conception of biological existence, which we hope may lead to important developments in the physics of life. The contextual nature of social constraints opens a fundamental question: *who does the constraint exist for?* Surely, if a society is organized around the activity of literate clerk caste, or a masculine warrior class, the specific constraints underlying the behavior of those elites will have important intensions in social organization. This is regardless of whether the general population is able to read and write, or to understand the locally afferent codification of masculine honor. In other words, a given cultural context may embed a coherent, closed system of constraints even though any single constraints may apply only to a limited portions of agents, and conversely any single agent may only be able to engage with specific dimensions of this context. At face value, this isn't a major concern: after all, we should expect the organizational of life to be an emergent property. However, this implicitly raises the issue of the relation between structure and information in the definition of life. From the perspective of an encultured agent, a social constraint is a prior belief (implicit or explicit) which shapes their experience of the world. But from the perspective of the sociocultural lifeforms they participate in, it is a basic element of its constitutive structure. In other words, there is a fundamental duality between the *structure* of the lifeforms and the underlying dynamics defining what constitutes *information* from the perspective of its participants. If the argument outlined in [5.4] is correct, the fundamental ambiguity between the semiotics of the organism's functional architecture and the semiotics of their structural implementation may provide a mechanism for the formation of novel semiotics, and hence for the open-ended evolution of life.

[5.12] Overall, the framework of social constraints affords a rich view of the articulation between human agency and social organization. On one hand, social constraints can be viewed as constraints over behavior and attention, which are embedded into the cultural ecology, integrated through enculturation, and permanently evolve through the reconstructive nature of both embedding and integration. On the other hand, social constraints can be viewed as constitutive elements of emerging sociocultural lifeforms, whose activity permanently reconstruct its own functional architecture. This entails that the functional architecture of sociocultural lifeforms is (by construction) embedded in its constitutive agents' embodied phenomenology. However, the perspective on the nature of social constraints of the social organization and its constituents diverge radically. To the organization, the constraints are constitutive elements of its own architecture, which its integrated activity (hopefully) reconstructs ; to the agents, they are the basic structure underlying their embodied phenomenology. This dual perspective opens an avenue for the investigation of social phenomena, as well as open-ended evolution in the biosphere more generally. By the present account, human agency is not exerted *against* the background of prior social structure. On the contrary, humans exert their agency by participating in sociocultural lifeforms, thereby driving the reconstruction (whether conservative or creative) of social structure. For the same reasons exposed in part 4, the process of alignment between the top-down and bottom-up perspective offers the potential for the construction of novel systems of abstracted communication, and therefore for the open-ended evolution of the space in which the system is embedded. Furthermore, this entire process is dual to the *internalization* and *externalization* of social constraints by individual agents. This provides us with the perspective of building computational experiments detailing the process of historical change through a model of cognitive engagement within a system of social constraints, and to do so with a far greater degree of precision that what is possible with narrative theory.

6) Conclusion

[6.1] We have hereby presented a novel framework for the description of human encultured cognition and social organization, which centers on the definition of "social constraints". In this context, social constraints refer to socially and materially embedded patterns in collective behavior, which encultured agents come to expect of each other and enact. A core element of this account stems from the dissolution of agency under Active Inference, or *embedded normativity* (Guénin--Carlut and Albarracin 2024). Agents under Active Inference are drawn to enact the very actions they expect of themselves and to sample the external world to extract the minimal sensory information they need to adequately regulate their behavior. Hence, they find their own activity canalized by the collection of norms (implicit or explicit) that are adequately embedded in their social and material niche. To illustrate this perspective, we may think of a hiker approaching an forest path. They may agree or disagree with the adequacy of the trajectory followed by the path, but they will have to account for the fact that deviating from it would require active effort. Overall, they are most likely to take the path of least effort and follow it without a second thought. We suggest that this attraction toward expected states or trajectories constitutes a straightforward expression of the fundamental physical processes underlying cognition. Following the mathematical framework of Bayesian gauge theory, we suggest that an agent's (implicit or explicit) expectations about their sensorimotor states translate into the emergence of (gauge) forces drawing their attention (and the implicit sensorimotor regulation it entails) toward the specific environmental cues which enable the realization of those expectations. This model of cognition explains many aspects of encultured cognition, assuming that agents are able and willing to abstract their own actions and sensations from their direct embodied experience and engage in what we have defined as the *existential stance*.

[6.2] Indeed, the ability to reflexively contextualize one's own activity as communicative, socially embedded acts enables the development of abstract modes of social regulation and their integration in one's direct engagement with the world. To walk a path to the river is to walk a path to the river, but to do so while anticipating how fetching water will help Grandma cook fundamentally alters the cognitive and emotional landscape of the activity. The meaning of walking then becomes less about path finding and balance, and more about reproducing early memories of eating and finding social validation in one's long term co-operative relation with Grandma. We do not mean to commit here to a specific evolutionary trajectory for the existential stance, but it seems clear that it may have developed as a byproduct of earlier modes of socially embedded emotional regulation such as dancing or musicality. Our main argument is that once it has emerged, it enables the development of a wide array of complex socially embedded practices, and of full-fledged encultured cognition as it can be found in contemporary humans. Throughout this article, we have developed the example of executive function (and working memory more specifically) as an *a priori* intrinsic cognitive function whose development and evolution can in essence be explained by the social integration of cognition through the existential stance. However, the most straightforward instance of that logic would be language. Through language, agents may negotiate public norms, beliefs, and intentions of arbitrary complexity through the (syntactic) composition of simple propositions articulated in well-defined semantics. Then, the same public commitments may be turned inwards to regulate one's mental activity, with the social embedding of linguistic interaction providing further motivation to actually monitor one's activity against one's words. This is, assuming agents agree on the value of specific utterances or signs as communicative act, embedding with specific meaning.

[6.3] However, the importance of the existential stance in shaping social organization is nowhere more critical than in the manner it enables agents to monitor their cognitive engagement with the world against abstracted social identities. The ability to communicate in an abstract manner about one's intentions and social relations affords the development of an abstracted sense of one's social role, embedded with norms of behavior. As a specific case of this process, intentional teaching is of special relevance to the early evolution of human encultured cognition. Indeed, it enables a student and a teacher co-operate in the transfer of specific skills from the teacher to the student. While no plausible physical mechanism may explain a direct transfer of embodied skill between different agents, the feedback offered by the teacher may help guide the student's attention toward the key parameters that mediate success in the relevant activity. This is, assuming both agents share the intention to transfer that skill, and do by assuming respectively the social role of teacher and student. Many other kinds of social roles have emerged through human history and became embedded in social organization. We could cite family roles such as parent and child, gender roles such as man and woman, ethnic identities, or professional/class roles such as warrior, priest or potter. Bayesian gauge theory and Active Inference offer an immediate explanation of how such roles, once developed, may become embedded in human cognition. Indeed, under Active Inference, one's perception of and engagement with the world stems from the continuous anticipation of one's sensations and actions. Once an encultured agent perceives their own activity through a given abstracted role, their anticipations become embedded in the social scripts, norms and expectations associated to that role. Hence, the drive toward Bayesian coherence push them toward the enaction of those specific roles, in the manner described by Bayesian gauge theory.

[6.4] The limitations of barebones Bayesian gauge theory appears when we consider the contextual nature of cognitive semiotics. In Bayesian gauge theory, the space of things an agent can perceive, the embodied expectations associated to those things, and the forces driving the agent toward the realization of those expectations stem from the *a priori* definition of the agent-environment system as a dynamical system. Respectively, they are defined as the space of external states, the information geometry stemming from the agent-environment partition, and the gauge force stemming from the symmetry between states of equal prior probability. In other words, Bayesian gauge theory assumes the existence of a well-defined space of external things to be perceived, and models cognitive engagement as a simple process of defining probabilistic belief over this space. By contrast, a god such as "Sol Invictus" or an institution such as "the Roman Empire" is not (like the mass, electrical charge, or chemical potential of an object) an objective state of the world which exists in isolation from its observation, and which may or may not be accurately assessed by a given observer. Its existence is entirely mediated by the shared belief in its existence by a community of encultured agents, and on the collection of norms that is associated to that belief. The same applies, in a lesser but significant extent, to perceptual concepts with an existing referent such as "to cook" or "Grandma". While those concepts indeed refer to something in the world, the specific manner in which they are operationalized (their cognitive meaning) constitutes the outcome of a constructive process where the agent (or communities thereof) decide on how they are to be defined and engaged with. We call "contextual" such properties, which are defined as an outcome of observation and in a manner contingent on the modalities of observation. Critically, contextual properties cannot be grounded in the bare dynamics of the agent-world system (see Guénin--Carlut (2024) regarding the relevance of this argument to cognition in general and Active Inference in particular).

[6.5] Therefore, explaining the construction of cultural landscapes of affordances requires another concept, that it quite like dynamics in the manner it constrains cognitive engagement, but quite unlike dynamics in the manner it is itself understood as the product of a contingent, constructive process. This concept is the keystone of our present account: *social constraints*. Our understanding of social constraints stems from the duality between cognitive dynamics and cognitive meaning outlined by Bayesian gauge theory, although re-operationalized in an adequate manner to account for the contextuality and open-ended evolution of cultural semiotics. Social constraints are understood to be the causal force underlying the contextual perception of the world through specific cultural concepts, in the sense that they canalize attention toward the adequate states of the world under the adequate norms to bring about those specific perceptive states. But at the same time, they flow from the structure of the world as perceived by the agent, in the manner prescribed by Bayesian gauge theory. Hence, they can account for the constructive and open-ended nature of cognitive engagement. Assume a naive existential agent engages with a cultural niche including the linguistic concept of "Grandma". Learning the meaning of that concept induces the integration of novel norms for cognitive engagement, which underlie what is to be perceived as "Grandma" or not-"Grandma" and how to engage with "Grandma". But once these norms are integrated, engaging with specific objects as instances of "Grandma" induces a reconfiguration of the cognitive landscape of the agent, who may for example learn to expect social and material support from "Grandma", as well as expect themselves to co-operate with "Grandma" along specific cultural scripts. The integration of social constraints bring about novel perceptive states, while novel perceptive states bring about novel social constraints. This grounds an open-ended cycle by which agents may develop cultural concepts embedded in their contextual experience of the world.

[6.6] Hence, the basic concept of social constraints as operationalized here induces a novel ontology for the social and cognitive sciences. It brings with it an array of novel methodological possibilities, as well as novel challenges for the study of human mind and culture. The meaning of this framework relies critically on the specific process by which agents may integrate social constraints, as well as the relation between the macro scale (this of the cultural landscape collectively experienced by cultural agents) and the micro scale (this of the dynamics of individual agents). Indeed, to exist in a meaningful sense, social constraints need to be embedded in the cognition of multiple different agent, whose physiological properties, social situations or developmental history may vary drastically. However, the process by which agents may integrate those constraints is inherently reconstructive, as no physical mechanism may plausibly "copy" the underlying mechanisms from mind to mind. Hence, the coherence of social constraints requires by construction that the concept applies to specific invariants of cognitive engagement that may effectively be robustly reconstructed by a variety of agents through engagement with a given material and cultural niche. The symbolic value of golden thrones, the topology of a urban system, or the meaning of "Grandma" may be described as a "social constraint", but the idiosyncratic manner in which specific agent engage with those concepts may not (unless it defines novel invariants which may be reproduced by peers). The functional account of constraints given by Montévil and Mossio (2015) is therefore a critical element of our framework: social constraints are themselves contextual objects, brought about by encultured perception defined by the prior social constraints agents have integrated. Hence, social constraints do not exist in and of themselves, they exist in the context of a broader system of constraints which is collectively able to (re)construct itself in the perception of cognitive agent through the process of enculturation.

[6.7] Regardless of the questions it opens, the most direct contribution of the framework to the methodology of social and cognitive science is most likely in the manner it frames the open debate on the relation between structure and agency. Our framework describes agency as a process expressed through the engagement with and the enaction of prior social structure – understood here as a system of social constraints. For example, I can only reach my goal of "helping Grandma cook" because I have developed a socially embedded understanding of what this notion means, along with socially embedded scripts and skills regarding how I can do it. In turn, the specific manner in which agents understand and enact social structure leads to its reconstruction in historical time. In itself, this notion is not novel, as it is central to both (Bourdieu 1972)'s practice theory and Giddens (1984)'s structuration theory. Our contribution rather stands in our specific model of the role of social structure in cognition. Rather than being simply a frame with which agents reflexively understand social relations, social structure (in the form of social constraints) participate in constituting one's basic experience of the world, in a way that scaffolds basic cognitive abilities. In this framework, we can relate engagement in social relations to the cultural evolution of cognitive abilities, as outlined in Heyes (2018)'s cultural evolutionary psychology.

Furthermore, we can study the manner in which specific technological practices or material landscapes direct the attention of the agent, so as to construct skill transfer or the local ideology underlying social relations. Finally, and most importantly, we may study the articulation between different scales of agency, as systems of social constraints inherit from the general properties of biological organization, and hence have the capacity to deploy their own norms and inscribe it in the material and social environment experienced by individual agent (Guénin--Carlut 2022e; Guénin--Carlut, White, and Sganzerla 2023b).

[6.8] To be clear, we do not claim to present here a fully articulated methodology to study the human mind and culture. In the absence of any consensus regarding the appropriate methodology for explaining human social activity (or even regarding the specific target of this inquiry), no single avenue seem to be able to address the many challenges opened by such an endeavor. However, we do propose a specific ontology for the evolution of human cognition and culture, which is fully grounded in and coherent with known elements of the physics of cognition. Of course, many elements are yet to be filled in. An explicit mathematical account of contextuality, or open-ended evolution, is not a trivial task (Kauffman and Roli 2023). Yet, we are inclined to believe that even as it stands here, our work has many interesting consequences for the study of mind and culture. It provides a clear account of the articulation between scales in social activity, which is coherent with (but more specific than) earlier authoritative treatments of the articulation between social structure and individual agency. It provides a meaningful description of how specific social norms and elements of the material niche shape human understanding of and engagement with their world. Perhaps most importantly, it provides clear constraints on possible scenarios for the evolution of human cognition and social organization. To our knowledge, this is the first framework which starts from the basic physical processes underlying cognition to ground the study of social organization and cultural evolution, in a manner that explicitly accounts for the contextuality of cultural concepts and/or the open-ended evolution of human culture. Therefore, it should demonstrate a privileged avenue for the study of cognition and culture, which dissolves by construction many of the methodological and conceptual difficulties plaguing XXth century sociology and cognitive science.

Acknowledgements

This work was financed by the XSCAPE project (Synergy Grant ERC-2020-SyG 951631). I would like to thank Iona Brenac, Yseult Héjja-Brichard, Alexis Rozanski, Maxime Trimolet, Daniel Friedman, and Andy Clark for their proofreading, as well as Manon Job for our discussions regarding the definition and relevance of this project.

Declaration of Interests

The author declares no competing interests

Bibliography

Albarracin, Mahault, Axel Constant, Karl Friston, and Maxwell James Ramstead. 2020. "A Variational Approach to Scripts." PsyArXiv. https://doi.org/10.31234/osf.io/67zy4.

Archer, Margaret Scotford. 2003. *Structure, Agency and the Internal Conversation*. Cambridge University Press.

Badcock, Paul B., Christopher G. Davey, Sarah Whittle, Nicholas B. Allen, and Karl Friston. 2017. "The Depressed Brain: An Evolutionary Systems Theory." *Trends in Cognitive Sciences* 21 (3): 182–94. https://doi.org/10.1016/j.tics.2017.01.005.

Baddeley, Alan. 2003. "Working Memory and Language: An Overview." *Journal of Communication Disorders*, ASHA 2002, 36 (3): 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4.

Baez, John. 2020. "Noether." February 17, 2020. https://math.ucr.edu/home/baez/noether.html.

Barkow, Jerome H., Leda Cosmides, and John Tooby. 1995. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford University Press.

Bateson, Patrick, and Kevin N. Laland. 2013. "Tinbergen's Four Questions: An Appreciation and an Update." *Trends in Ecology & Evolution* 28 (12): 712–18. https://doi.org/10.1016/j.tree.2013.09.013.

Baudrillard, Jean. 1994. *Simulacra and Simulation*. University of Michigan Press.

Bechtel, William. 2008. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. Taylor & Francis.

———. 2009. "Looking down, around, and up: Mechanistic Explanation in Psychology." *Philosophical Psychology* 22 (5): 543–64. https://doi.org/10.1080/09515080903238948.

Bechtel, William, and Adele Abrahamsen. 2005. "Explanation: A Mechanist Alternative." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, Mechanisms in biology, 36 (2): 421–41. https://doi.org/10.1016/j.shpsc.2005.03.010.

Boone, Worth, and Gualtiero Piccinini. 2016. "The Cognitive Neuroscience Revolution." *Synthese* 193 (5): 1509–34. https://doi.org/10.1007/s11229-015-0783-4.

Bourdieu, Pierre. 1972. "Esquisse d'une Théorie de La Pratique [Outline of a Theory of Practice]." *Paris: Éditions Droz*.

Boyd, Robert, Peter J. Richerson, and Joseph Henrich. 2011. "The Cultural Niche: Why Social Learning Is Essential for Human Adaptation." *Proceedings of the National Academy of Sciences* 108 (Supplement 2): 10918–25. https://doi.org/10.1073/pnas.1100290108.

Brooks, Rodney A. 1991. "Intelligence without Representation." *Artificial Intelligence* 47 (1): 139–59. https://doi.org/10.1016/0004-3702(91)90053-M.

Bruineberg, Jelle, and Erik Rietveld. 2014. "Self-Organization, Free Energy Minimization, and Optimal Grip on a Field of Affordances." *Frontiers in Human Neuroscience* 8. https://doi.org/10.3389/fnhum.2014.00599.

Chatwin, Bruce. 2012. *The Songlines*. Random House.

Claidière, Nicolas, Thomas C. Scott-Phillips, and Dan Sperber. 2014. "How Darwinian Is Cultural Evolution?" *Philosophical Transactions of the Royal Society B: Biological Sciences* 369 (1642): 20130368. https://doi.org/10.1098/rstb.2013.0368.

Claidière, Nicolas, and Dan Sperber. 2007. “The Role of Attraction in Cultural Evolution.” *Journal of Cognition and Culture* 7 (1–2): 89–111. https://doi.org/10.1163/156853707X171829.
Clark, Andy. 2006. “Language, Embodiment, and the Cognitive Niche.” *Trends in Cognitive Sciences* 10 (8): 370–74. https://doi.org/10.1016/j.tics.2006.06.012.
———. 2013. “Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science.” *Behavioral and Brain Sciences* 36 (3): 181–204. https://doi.org/10.1017/S0140525X12000477.
———. 2015. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press.
———. 2020. “Beyond Desire? Agency, Choice, and the Predictive Mind.” *Australasian Journal of Philosophy* 98 (1): 1–15. https://doi.org/10.1080/00048402.2019.1602661.
Comte, Auguste. 1844. *Discours sur l’esprit positif*. Carilian-Goeury et V. Dalmont.
Constant, Axel, Andy Clark, and Karl Friston. 2021. “Representation Wars: Enacting an Armistice Through Active Inference.” *Frontiers in Psychology* 11:3798. https://doi.org/10.3389/fpsyg.2020.598733.
Coolidge, Frederick L., and Thomas Wynn. 2005. “Working Memory, Its Executive Functions, and the Emergence of Modern Thinking.” *Cambridge Archaeological Journal* 15 (1): 5–26. https://doi.org/10.1017/S0959774305000016.
Corcoran, Andrew W., and Jakob Hohwy. 2017. “Allostasis, Interoception, and the Free Energy Principle: Feeling Our Way Forward.” PsyArXiv. https://doi.org/10.31234/osf.io/zbqnx.
Cosmides, Leda, and John Tooby. 1992. “Cognitive Adaptations for Social Exchange.” *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* 163:163–228.
———. 2007. *Evolutionary Psychology: A Primer*.
Cristofori, Irene, Shira Cohen-Zimerman, and Jordan Grafman. 2019. “Chapter 11 - Executive Functions.” In *Handbook of Clinical Neurology*, edited by Mark D’Esposito and Jordan H. Grafman, 163:197–219. The Frontal Lobes. Elsevier. https://doi.org/10.1016/B978-0-12-804281-6.00011-2.
Da Costa, Lancelot, Karl Friston, Conor Heins, and Grigorios A. Pavliotis. 2021. “Bayesian Mechanics for Stationary Processes.” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 477 (2256): 20210518. https://doi.org/10.1098/rspa.2021.0518.
Di Paolo, Ezequiel. 2005. “Autopoiesis, Adaptivity, Teleology, Agency.” *Phenomenology and the Cognitive Sciences* 4 (4): 429–52. https://doi.org/10.1007/s11097-005-9002-y.
Diamond, Adele. 2013. “Executive Functions.” *Annual Review of Psychology* 64:135–68. https://doi.org/10.1146/annurev-psych-113011-143750.
Duhem, Pierre. 2016. *La théorie physique. Son objet, sa structure*. ENS Éditions.
Durkheim, Emile. 1982. “Rules for the Explanation of Social Facts.” In *The Rules of Sociological Method: And Selected Texts on Sociology and Its Method*, edited by Steven Lukes, 119–46. Contemporary Social Theory. London: Macmillan Education UK. https://doi.org/10.1007/978-1-349-16939-9_6.
Fodor, Jerry A. 1975. *The Language of Thought*. Harvard University Press.
Fodor, Jerry, and Zenon W. Pylyshyn. 1988. “Connectionnism and Cognitive Architecture : A Critical Analysis,” 1988.
Foucault, Michel. 1980. “Truth and Power In Gordon, C.(Ed.) Power/Knowledge: Selected Interviews and Other Writings 1972-77.” Brighton: Harvester Wheatsheaf.

Friston, Karl. 2010. "The Free-Energy Principle: A Unified Brain Theory?" *Nature Reviews Neuroscience* 11 (2): 127–38. https://doi.org/10.1038/nrn2787.

———. 2012. "A Free Energy Principle for Biological Systems." *Entropy* 14 (11): 2100–2121. https://doi.org/10.3390/e14112100.

———. 2013. "Life as We Know It." *Journal of The Royal Society Interface* 10 (86): 20130475. https://doi.org/10.1098/rsif.2013.0475.

———. 2019. "A Free Energy Principle for a Particular Physics." *arXiv:1906.10184 [q-Bio]*, June. http://arxiv.org/abs/1906.10184.

Friston, Karl, Lancelot Da Costa, Danijar Hafner, Casper Hesp, and Thomas Parr. 2021. "Sophisticated Inference." *Neural Computation* 33 (3): 713–63. https://doi.org/10.1162/neco_a_01351.

Friston, Karl, Lancelot Da Costa, Dalton A. R. Sakthivadivel, Conor Heins, Grigorios A. Pavliotis, Maxwell Ramstead, and Thomas Parr. 2023. "Path Integrals, Particular Kinds, and Strange Things." *Physics of Life Reviews*, August. https://doi.org/10.1016/j.plrev.2023.08.016.

Friston, Karl, Jean Daunizeau, and Stefan J. Kiebel. 2009. "Reinforcement Learning or Active Inference?" *PLOS ONE* 4 (7): e6421. https://doi.org/10.1371/journal.pone.0006421.

Friston, Karl, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, John O'Doherty, and Giovanni Pezzulo. 2016. "Active Inference and Learning." *Neuroscience & Biobehavioral Reviews* 68 (September):862–79. https://doi.org/10.1016/j.neubiorev.2016.06.022.

Friston, Karl, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, and Giovanni Pezzulo. 2016. "Active Inference: A Process Theory." *Neural Computation* 29 (1): 1–49. https://doi.org/10.1162/NECO_a_00912.

Friston, Karl, and Stefan Kiebel. 2009. "Predictive Coding under the Free-Energy Principle." *Philosophical Transactions of the Royal Society B: Biological Sciences* 364 (1521): 1211–21. https://doi.org/10.1098/rstb.2008.0300.

Friston, Karl, James Kilner, and Lee Harrison. 2006. "A Free Energy Principle for the Brain." *Journal of Physiology-Paris*, Theoretical and Computational Neuroscience: Understanding Brain Functions, 100 (1): 70–87. https://doi.org/10.1016/j.jphysparis.2006.10.001.

Friston, Karl, Thomas Parr, and Bert de Vries. 2017. "The Graphical Brain: Belief Propagation and Active Inference." *Network Neuroscience*. https://doi.org/10.1162/NETN_a_00018.

Froese, Tom. 2022. "Scientific Observation Is Socio-Materially Augmented Perception: Toward a Participatory Realism." *Philosophies* 7 (2): 37. https://doi.org/10.3390/philosophies7020037.

Fuchs, Christopher A. 2017. "On Participatory Realism." In *Information and Interaction: Eddington, Wheeler, and the Limits of Knowledge*, edited by Ian T. Durham and Dean Rickles, 113–34. The Frontiers Collection. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-43760-6_7.

Fulminante, Francesca, Luce Prignano, Ignacio Morer, and Sergi Lozano. 2017. "Coordinated Decisions and Unbalanced Power. How Latin Cities Shaped Their Terrestrial Transportation Network." *Frontiers in Digital Humanities* 4. https://doi.org/10.3389/fdigh.2017.00004.

Giddens, Anthony. 1984. *The Constitution of Society: Outline of the Theory of Structuration*. University of California Press.

Goffman, Erving. 2022. *The Presentation of Self in Everyday Life*. Penguin Books, Limited.

Green, Sara. 2015. “Revisiting Generality in Biology: Systems Biology and the Quest for Design Principles.” *Biology & Philosophy* 30 (5): 629–52. https://doi.org/10.1007/s10539-015-9496-9.
Green, Sara, and Nicholaos Jones. 2016. “Constraint-Based Reasoning for Search and Explanation: Strategies for Understanding Variation and Patterns in Biology.” *Dialectica* 70 (3): 343–74. https://doi.org/10.1111/1746-8361.12145.
Guénin--Carlut, Avel. 2022a. “Cognitive Agency in Sociocultural Evolution.” OSF Preprints. https://doi.org/10.31219/osf.io/x7yr4.
———. 2022b. “On Participatory Realism.” *Kairos Journal*, October 18, 2022. https://kairos-research.org/on-participatory-realism.
———. 2022c. “Physics of Creation - Symmetry Breaking, (En)Active Inference, and Unfolding Statespaces.” OSF Preprints. https://doi.org/10.31219/osf.io/68947.
———. 2022d. “The Clothes of the Empire : An Active Inference Account of Identity Capture.” *Kairos Journal*, May 2022. https://kairos-research.org/clothes-of-the-empire.
———. 2022e. “Thinking like a State: Embodied Intelligence in the Deep History of Our Collective Mind.” *IOP Conference Series: Materials Science and Engineering* 1261 (1): 012026. https://doi.org/10.1088/1757-899X/1261/1/012026.
———. 2023. “Strange Things, Statespace Representation, and Participatory Realism - Comment on ‘Path Integrals, Particular Kinds, and Strange Things.’ by Friston et Al.” *Physics of Life Reviews*, October. https://doi.org/10.1016/j.plrev.2023.10.027.
———. 2024. “Contextuality, Cognitive Engagement, and Active Inference.” OSF. https://doi.org/10.31219/osf.io/wtrgh.
Guénin--Carlut, Avel, and Mahault Albarracin. 2024. “On Embedded Normativity an Active Inference Account of Agency Beyond Flesh.” In *Active Inference*, edited by Christopher L. Buckley, Daniela Cialfi, Pablo Lanillos, Maxwell Ramstead, Noor Sajid, Hideaki Shimazaki, Tim Verbelen, and Martijn Wisse, 91–105. Communications in Computer and Information Science. Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-47958-8_7.
Guénin--Carlut, Avel, and Iona Brenac. 2023. “Social Constraints - Graphic,” March. https://osf.io/chwxm/.
Guénin--Carlut, Avel, Ben White, and Lorena Sganzerla. 2023a. “The Cognitive Archaeology of Sociocultural Lifeforms.” In *ALIFE 2023: Ghost in the Machine: Proceedings of the 2023 Artificial Life Conference*. MIT Press. https://doi.org/10.1162/isal_a_00670.
———. 2023b. “The Cognitive Archeology of Sociocultural Lifeforms.” OSF Preprints. https://doi.org/10.31219/osf.io/qxszh.
Haan, Sanneke de. 2017. “The Existential Dimension in Psychiatry: An Enactive Framework.” *Mental Health, Religion & Culture* 20 (6): 528–35. https://doi.org/10.1080/13674676.2017.1378326.
Henrich, Joseph. 2017. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.
Heyes, Cecilia. 2012a. “Grist and Mills: On the Cultural Origins of Cultural Learning.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1599): 2181–91. https://doi.org/10.1098/rstb.2012.0120.
———. 2012b. “What’s Social about Social Learning?” *Journal of Comparative Psychology* 126 (2): 193–202. https://doi.org/10.1037/a0025180.
———. 2016. “Who Knows? Metacognitive Social Learning Strategies.” *Trends in Cognitive Sciences* 20 (3): 204–13. https://doi.org/10.1016/j.tics.2015.12.007.

———. 2018. *Cognitive Gadgets: The Cultural Evolution of Thinking*. Harvard University Press.
———. 2020. “Culture Primer.” Preprint. PsyArXiv. https://doi.org/10.31234/osf.io/v8ms3.
———. 2023. “Rethinking Norm Psychology.” *Perspectives on Psychological Science*, July, 17456916221112075. https://doi.org/10.1177/17456916221112075.
Heyes, Cecilia, and Caroline Catmur. 2022. “What Happened to Mirror Neurons?” *Perspectives on Psychological Science* 17 (1): 153–68. https://doi.org/10.1177/1745691621990638.
Hipólito, Inês, Manuel Baltieri, Karl Friston, and Maxwell James Ramstead. 2021. “Embodied Skillful Performance: Where the Action Is.” *Synthese*, January. https://doi.org/10.1007/s11229-020-02986-5.
Hohwy, Jakob. 2016. “The Self-Evidencing Brain.” *Noûs* 50 (2): 259–85. https://doi.org/10.1111/nous.12062.
Hooft, Gerard ′t. 2008. “Gauge Theories.” *Scholarpedia* 3 (12): 7443. https://doi.org/10.4249/scholarpedia.7443.
Hosoya, Toshihiko, Stephen A. Baccus, and Markus Meister. 2005. “Dynamic Predictive Coding by the Retina.” *Nature* 436 (7047): 71–77. https://doi.org/10.1038/nature03689.
Jablonka, Eva. 1994. “Inheritance Systems and the Evolution of New Levels of Individuality.” *Journal of Theoretical Biology* 170 (3): 301–9. https://doi.org/10.1006/jtbi.1994.1191.
———. 2002. “Information: Its Interpretation, Its Inheritance, and Its Sharing.” *Philosophy of Science* 69 (4): 578–605. https://doi.org/10.1086/344621.
Jablonka, Eva, and Marion J. Lamb. 2006. “The Evolution of Information in the Major Transitions.” *Journal of Theoretical Biology*, Special Issue in Memory of John Maynard Smith, 239 (2): 236–46. https://doi.org/10.1016/j.jtbi.2005.08.038.
———. 2008. “Soft Inheritance: Challenging the Modern Synthesis.” *Genetics and Molecular Biology* 31 (2): 389–95. https://doi.org/10.1590/S1415-47572008000300001.
Kauffman, Stuart A. 1969. “Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets.” *Journal of Theoretical Biology* 22 (3): 437–67. https://doi.org/10.1016/0022-5193(69)90015-0.
———. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press.
———. 2000. *Investigations*. Oxford University Press.
———. 2014. “Prolegomenon to Patterns in Evolution.” *Biosystems*, SI :Patterns in Evolution, 123 (September):3–8. https://doi.org/10.1016/j.biosystems.2014.03.004.
Kauffman, Stuart A., and Andrea Roli. 2023. “A Third Transition in Science?” *Interface Focus* 13 (3): 20220063. https://doi.org/10.1098/rsfs.2022.0063.
Kirchhoff, Michael D. 2018. “Autopoiesis, Free Energy, and the Life–Mind Continuity Thesis.” *Synthese* 195 (6): 2519–40. https://doi.org/10.1007/s11229-016-1100-6.
Kirchhoff, Michael, Thomas Parr, Ensor Palacios, Karl Friston, and Julian Kiverstein. 2018. “The Markov Blankets of Life: Autonomy, Active Inference and the Free Energy Principle.” *Journal of The Royal Society Interface* 15 (138): 20170792. https://doi.org/10.1098/rsif.2017.0792.
Kocagoncu, Ece, Anastasia Klimovich-Gray, Laura E Hughes, and James B Rowe. 2021. “Evidence and Implications of Abnormal Predictive Coding in Dementia.” *Brain* 144 (11): 3311–21. https://doi.org/10.1093/brain/awab254.

Laland, K. N., J. Odling-Smee, and M. W. Feldman. 2001. “Cultural Niche Construction and Human Evolution.” *Journal of Evolutionary Biology* 14 (1): 22–33. https://doi.org/10.1046/j.1420-9101.2001.00262.x.

Laland, Kevin N., Tobias Uller, Marcus W. Feldman, Kim Sterelny, Gerd B. Müller, Armin Moczek, Eva Jablonka, and John Odling-Smee. 2015. “The Extended Evolutionary Synthesis: Its Structure, Assumptions and Predictions.” *Proc. R. Soc. B* 282 (1813): 20151019. https://doi.org/10.1098/rspb.2015.1019.

Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. OUP Oxford.

Lent, Jeremy. 2017. *The Patterning Instinct: A Cultural History of Humanity’s Search for Meaning*. Prometheus Books.

Lévi-Strauss, Claude. 1958. *Anthropologie structurale: avec 13 shémas dans le texte*. Plon.

Longo, Giuseppe, Maël Montévil, and Stuart Kauffman. 2012. “No Entailing Laws, but Enablement in the Evolution of the Biosphere.” In *Proceedings of the 14th Annual Conference Companion on Genetic and Evolutionary Computation*, 1379–92. GECCO ’12. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2330784.2330946.

Maturana, H. R., and F. J. Varela. 1991. *Autopoiesis and Cognition: The Realization of the Living*. Springer Science & Business Media.

Mercier, Hugo, and Dan Sperber. 2011. “Why Do Humans Reason? Arguments for an Argumentative Theory.” *Behavioral and Brain Sciences* 34 (2): 57–74; discussion 74-111. https://doi.org/10.1017/S0140525X10000968.

Mitchell, Sandra D. 2003. *Biological Complexity and Integrative Pluralism*. Cambridge University Press.

Montévil, Maël, and Matteo Mossio. 2015. “Biological Organisation as Closure of Constraints.” *Journal of Theoretical Biology* 372 (May):179–91. https://doi.org/10.1016/j.jtbi.2015.02.029.

Moreno, Alvaro, and Matteo Mossio. 2015. “Biological Autonomy. A Philosophical and Theoretical Enquiry,” May. https://www.academia.edu/12363315/Biological_Autonomy_A_Philosophical_and_Theoretical_Enquiry.

Mulder, Monique Borgerhoff, Richard McElreath, and Kari Britt Schroeder. 2006. “Analogies Are Powerful and Dangerous Things.” *Behavioral and Brain Sciences* 29 (4): 350–51. https://doi.org/10.1017/S0140525X06259089.

Noether, Emmy. 1971. “Invariant Variation Problems.” *Transport Theory and Statistical Physics* 1 (3): 186–207. https://doi.org/10.1080/00411457108231446.

Norris, Ray P., and Bill Yidumduma Harney. 2014. “Songlines and Navigation in Wardaman and Other Australian Aboriginal Cultures.” arXiv. https://doi.org/10.48550/arXiv.1404.2361.

Ogien, Albert. 2011. “Les sciences cognitives ne sont pas des sciences humaines. Une réponse à « Vers un naturalisme social » de Laurence Kaufman et Laurent Cordonier.” *SociologieS*, October. http://journals.openedition.org/sociologies/3635.

O’Raifeartaigh, Lochlainn, and Norbert Straumann. 2000. “Gauge Theory: Historical Origins and Some Modern Developments.” *Reviews of Modern Physics* 72 (1): 1–23. https://doi.org/10.1103/RevModPhys.72.1.

Parr, Thomas, Lancelot Da Costa, and Karl Friston. 2020. “Markov Blankets, Information Geometry and Stochastic Thermodynamics.” *Philosophical Transactions of the Royal*

Society A: Mathematical, Physical and Engineering Sciences 378 (2164): 20190159. https://doi.org/10.1098/rsta.2019.0159.
Parr, Thomas, Giovanni Pezzulo, and Karl Friston. 2022. *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. Cambridge, MA, USA: MIT Press.
Piccinini, Gualtiero. 2010. “The Mind as Neural Software? Understanding Functionalism, Computationalism, and Computational Functionalism.” *Philosophy and Phenomenological Research* 81 (2): 269–311. https://doi.org/10.1111/j.1933-1592.2010.00356.x.
Pigliucci, Massimo, and Gerd Muller. 2010. *Evolution – the Extended Synthesis*. MIT Press.
Pika, Simone, Miriam Jennifer Sima, Christian R. Blum, Esther Herrmann, and Roger Mundry. 2020. “Ravens Parallel Great Apes in Physical and Social Cognitive Skills.” *Scientific Reports* 10 (1): 20617. https://doi.org/10.1038/s41598-020-77060-8.
Prigogine, I., and G. Nicolis. 1971. “Biological Order, Structure and Instabilities.” *Quarterly Reviews of Biophysics* 4 (2–3): 107–48. https://doi.org/10.1017/S0033583500000615.
Quine, W. V. 1951. “Main Trends in Recent Philosophy: Two Dogmas of Empiricism.” *The Philosophical Review* 60 (1): 20–43. https://doi.org/10.2307/2181906.
Ramstead, Maxwell James, Karl Friston, and Inês Hipólito. 2020. “Is the Free-Energy Principle a Formal Theory of Semantics? From Variational Density Dynamics to Neural and Phenotypic Representations.” *Entropy* 22 (8): 889. https://doi.org/10.3390/e22080889.
Ramstead, Maxwell James, Michael D Kirchhoff, and Karl Friston. 2020. “A Tale of Two Densities: Active Inference Is Enactive Inference.” *Adaptive Behavior* 28 (4): 225–39. https://doi.org/10.1177/1059712319862774.
Ramstead, Maxwell James, Dalton A. R. Sakthivadivel, Conor Heins, Magnus Koudahl, Beren Millidge, Lancelot Da Costa, Brennan Klein, and Karl Friston. 2022. “On Bayesian Mechanics: A Physics of and by Beliefs.” arXiv:2205.11543. arXiv. https://doi.org/10.48550/arXiv.2205.11543.
Ramstead, Maxwell, Samuel PL Veissière, and Laurence J. Kirmayer. 2016. “Cultural Affordances: Scaffolding Local Worlds through Shared Intentionality and Regimes of Attention.” *Frontiers in Psychology* 7:1090.
Rivers, Jacqueline J., and Robert A. Josephs. 2010. “Dominance and Health: The Role of Social Rank in Physiology and Illness.” In *The Social Psychology of Power*, 87–112. New York, NY, US: The Guilford Press.
Ross, Lauren N. 2023. “The Explanatory Nature of Constraints: Law-Based, Mathematical, and Causal.” *Synthese* 202 (2): 56. https://doi.org/10.1007/s11229-023-04281-5.
Ruiz-Mirazo, Kepa, and Alvaro Moreno. 2004. “Basic Autonomy as a Fundamental Step in the Synthesis of Life.” *Artificial Life* 10 (3): 235–59. https://doi.org/10.1162/1064546041255584.
Ruiz-Mirazo, Kepa, Juli Peretó, and Alvaro Moreno. 2004. “A Universal Definition of Life: Autonomy and Open-Ended Evolution.” *Origins of Life and Evolution of the Biosphere* 34 (3): 323–46. https://doi.org/10.1023/B:ORIG.0000016440.53346.dc.
Sakthivadivel, Dalton A. R. 2022. “A Constraint Geometry for Inference and Integration.” *arXiv:2203.08119 [Cond-Mat, Physics:Math-Ph]*, March. http://arxiv.org/abs/2203.08119.
Sasaki, Takao, and Dora Biro. 2017. “Cumulative Culture Can Emerge from Collective Intelligence in Animal Groups.” *Nature Communications* 8 (1): 15049. https://doi.org/10.1038/ncomms15049.

Savage, Patrick E., Psyche Loui, Bronwyn Tarr, Adena Schachner, Luke Glowacki, Steven Mithen, and W. Tecumseh Fitch. 2021a. “Music as a Coevolved System for Social Bonding.” *Behavioral and Brain Sciences* 44:1–36. https://doi.org/10.1017/S0140525X20000333.

———. 2021b. “Toward Inclusive Theories of the Evolution of Musicality.” *Behavioral and Brain Sciences* 44. https://doi.org/10.1017/S0140525X21000042.

Schrodinger, Erwin. 1944. “What Is Life.”

Scott, James C. 2017. *Against the Grain: A Deep History of the Earliest States*. Yale University Press.

Sengupta, Biswa, and Karl Friston. 2016. “Approximate Bayesian Inference as a Gauge Theory.” *PLOS Biology* 14 (3): e1002400. https://doi.org/10.1371/journal.pbio.1002400.

Sengupta, Biswa, Arturo Tozzi, Gerald K. Cooray, Pamela K. Douglas, and Karl Friston. 2016. “Towards a Neuronal Gauge Theory.” *PLOS Biology* 14 (3): e1002400. https://doi.org/10.1371/journal.pbio.1002400.

Seth, Anil K. 2013. “Interoceptive Inference, Emotion, and the Embodied Self.” *Trends in Cognitive Sciences* 17 (11): 565–73. https://doi.org/10.1016/j.tics.2013.09.007.

Shilton, Dor, Mati Breski, Daniel Dor, and Eva Jablonka. 2020. “Human Social Evolution: Self-Domestication or Self-Control?” *Frontiers in Psychology* 11. https://doi.org/10.3389/fpsyg.2020.00134.

Slingerland, Edward. 2014. *Trying Not to Try: The Ancient Art of Effortlessness and the Surprising Power of Spontaneity*. Canongate Books.

Soucy, Shannon M., Jinling Huang, and Johann Peter Gogarten. 2015. “Horizontal Gene Transfer: Building the Web of Life.” *Nature Reviews Genetics* 16 (8): 472–82. https://doi.org/10.1038/nrg3962.

Sperber, Dan. 1996. *Explaining Culture : A Naturalistic Approach*.

———. 2001. “In Defense of Massive Modularity.” *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler* 7:47–57.

———. 2011. “A Naturalistic Ontology for Mechanistic Explanations in the Social Sciences.” *Analytical Sociology and Social Mechanisms*, 64–77.

Sperber, Dan, and Lawrence A. Hirschfeld. 2004. “The Cognitive Foundations of Cultural Stability and Diversity,” 2004.

Sperber, Dan, and Hugo Mercier. 2012. “Reasoning as a Social Competence.” *Collective Wisdom: Principles and Mechanisms*, 368–92.

Sterelny, Kim. 2007. “Social Intelligence, Human Intelligence and Niche Construction.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 362 (1480): 719–30. https://doi.org/10.1098/rstb.2006.2006.

———. 2011. “From Hominins to Humans: How Sapiens Became Behaviourally Modern.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 366 (1566): 809–22. https://doi.org/10.1098/rstb.2010.0301.

Sterzer, Philipp, Rick A. Adams, Paul Fletcher, Chris Frith, Stephen M. Lawrie, Lars Muckli, Predrag Petrovic, Peter Uhlhaas, Martin Voss, and Philip R. Corlett. 2018. “The Predictive Coding Account of Psychosis.” *Biological Psychiatry*, Mechanisms of Cognitive Impairment in Schizophrenia, 84 (9): 634–43. https://doi.org/10.1016/j.biopsych.2018.05.015.

Stewart, John E. 2020. “Towards a General Theory of the Major Cooperative Evolutionary Transitions.” *Biosystems* 198 (December):104237. https://doi.org/10.1016/j.biosystems.2020.104237.

Stout, Dietrich. 2021. “The Cognitive Science of Technology.” *Trends in Cognitive Sciences* 25 (11): 964–77. https://doi.org/10.1016/j.tics.2021.07.005.
Sutton, John, Celia B. Harris, Paul G. Keil, and Amanda J. Barnier. 2010. “The Psychology of Memory, Extended Cognition, and Socially Distributed Remembering.” *Phenomenology and the Cognitive Sciences* 9 (4): 521–60. https://doi.org/10.1007/s11097-010-9182-y.
Szathmáry, Eörs. 2015. “Toward Major Evolutionary Transitions Theory 2.0.” *Proceedings of the National Academy of Sciences* 112 (33): 10104–11. https://doi.org/10.1073/pnas.1421398112.
Szathmáry, Eörs, and John Maynard Smith. 1995. “The Major Evolutionary Transitions.” *Nature* 374 (6519): 227. https://doi.org/10.1038/374227a0.
Tarasi, Luca, Jelena Trajkovic, Stefano Diciotti, Giuseppe di Pellegrino, Francesca Ferri, Mauro Ursino, and Vincenzo Romei. 2022. “Predictive Waves in the Autism-Schizophrenia Continuum: A Novel Biobehavioral Model.” *Neuroscience & Biobehavioral Reviews* 132 (January):1–22. https://doi.org/10.1016/j.neubiorev.2021.11.006.
Tenenbaum, Joshua B., Thomas L. Griffiths, and Charles Kemp. 2006. “Theory-Based Bayesian Models of Inductive Learning and Reasoning.” *Trends in Cognitive Sciences*, Special issue: Probabilistic models of cognition, 10 (7): 309–18. https://doi.org/10.1016/j.tics.2006.05.009.
Tenenbaum, Joshua B., Charles Kemp, Thomas L. Griffiths, and Noah D. Goodman. 2011. “How to Grow a Mind: Statistics, Structure, and Abstraction.” *Science* 331 (6022): 1279–85. https://doi.org/10.1126/science.1192788.
Tison, Remi. 2022. “The Fanciest Sort of Intentionality: Active Inference, Mindshaping and Linguistic Content.” *Philosophical Psychology* 0 (0): 1–41. https://doi.org/10.1080/09515089.2022.2062315.
Tison, Remi, and Pierre Poirier. 2021. “Communication as Socially Extended Active Inference: An Ecological Approach to Communicative Behavior.” *Ecological Psychology* 33 (3–4): 197–235. https://doi.org/10.1080/10407413.2021.1965480.
Tomasello, Michael. 2016. “The Ontogeny of Cultural Learning.” *Current Opinion in Psychology*, Culture, 8 (April):1–4. https://doi.org/10.1016/j.copsyc.2015.09.008.
Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll. 2005. “Understanding and Sharing Intentions: The Origins of Cultural Cognition.” *Behavioral and Brain Sciences* 28 (5): 675–91. https://doi.org/10.1017/S0140525X05000129.
Tomasello, Michael, and Henrike Moll. 2010. “The Gap Is Social: Human Shared Intentionality and Culture.” In *Mind the Gap: Tracing the Origins of Human Universals,* edited by Peter M. Kappeler and Joan Silk, 331–49. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-02725-3_16.
Varela, F. G., H. R. Maturana, and R. Uribe. 1974. “Autopoiesis: The Organization of Living Systems, Its Characterization and a Model.” *Biosystems* 5 (4): 187–96. https://doi.org/10.1016/0303-2647(74)90031-8.
Varela, Francisco J., Evan Thompson, and Eleanor Rosch. 2016. *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.
Veissière, Samuel P. L., Axel Constant, Maxwell James Ramstead, Karl Friston, and Laurence J. Kirmayer. 2020. “Thinking through Other Minds: A Variational Approach to Cognition and Culture.” *Behavioral and Brain Sciences* 43. https://doi.org/10.1017/S0140525X19001213.

Wagner, Andreas. 2008. “Robustness and Evolvability: A Paradox Resolved.” *Proceedings of the Royal Society B: Biological Sciences* 275 (1630): 91–100. https://doi.org/10.1098/rspb.2007.1137.

———. 2012. “The Role of Robustness in Phenotypic Adaptation and Innovation.” *Proceedings of the Royal Society B: Biological Sciences* 279 (1732): 1249–58. https://doi.org/10.1098/rspb.2011.2293.

Wagner, Günter P., and Jianzhi Zhang. 2011. “The Pleiotropic Structure of the Genotype–Phenotype Map: The Evolvability of Complex Organisms.” *Nature Reviews Genetics* 12 (3): 204–13.

Walsh, Kevin S., David P. McGovern, Andy Clark, and Redmond G. O’Connell. 2020. “Evaluating the Neurophysiological Evidence for Predictive Processing as a Model of Perception.” *Annals of the New York Academy of Sciences* 1464 (1): 242–68. https://doi.org/10.1111/nyas.14321.

Williams, Daniel. 2023. “The Marketplace of Rationalizations.” *Economics & Philosophy* 39 (1): 99–123. https://doi.org/10.1017/S0266267121000389.

Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. Macmillan.

Wrangham, Richard. 2017. “Control of Fire in the Paleolithic: Evaluating the Cooking Hypothesis.” *Current Anthropology* 58 (S16): S303–13. https://doi.org/10.1086/692113.