REVISTA OWL (*OWL Journal*)
www.revistaowl.com.br – ISSN: 2965-2634

# SO WHAT? GENERATIVE AI AND THE HISTORICAL REDEFINITION OF COGNITIVE COMPETENCE

## *E DAÍ? A IA GENERATIVA E A REDEFINIÇÃO HISTÓRICA DAS COMPETÊNCIAS COGNITIVAS*

DOI: 10.5281/zenodo.21520072

*Julio Cesar de Aguiar*[1]

**ABSTRACT**

Generative AI has entered intellectual practice with a quality best described as irresistible. Much of the intellectual response in research and education has been organized around an alarm: that these systems induce a surrender of thinking, an atrophy of cognitive capacity, an accumulation of cognitive debt. This essay argues that the alarm conflates two claims that must be told apart. The first is categorical — that generative AI corrupts a cognitive interior whose integrity consists in operating unaided. Its governing metaphors, debt and atrophy and renunciation, presuppose a prior solvent interior against which the deficit is scored, and no such interior is to be found at any documented moment of the historical record: the Greek transition from orality to literacy shows a comparable cognitive figure lost without loss of the human. The second claim is empirical and concerns the ecology in which competences are formed. It survives the separation, and the essay relocates rather than refutes it. Drawing on Latour's argument that we have never been modern, and on convergent descriptions of cognition as coupled practice (Hutchins, Stiegler, Malafouris, Barad, Goody, Ong, Yuk Hui), the essay reconstructs cognition as a configuration of brains, bodies, tools, environments, and other people — a configuration whose history has repeatedly redefined what counts as cognitive competence. Recent measures of cognitive loss under generative AI assess competences with the ruler of the configuration those competences are leaving behind. What survives the critique is a systemic question, left open here: whether the emerging configuration sustains the practice it draws upon or depletes it. The burden of showing that it does not falls on those who raise the alarm. Concerns that are political-economic, environmental, labor-related, and about opacity and cultural standardization survive examination and deserve deliberation. The categorical alarm about cognition does not.

1 Professor da Escola de Políticas Públicas, Governo e Empresas da Fundação Getulio Vargas (EPPG-FGV), Brasília, Brasil. E-mail: julio.aguiar@fgv.br. ORCID: https://orcid.org/0000-0002-8252-2894.

**Keywords:** Generative AI; Cognitive internalism; Cognitive offloading; Distributed cognition; Modern Constitution; Hybrid cognition.

**RESUMO**

A IA generativa entrou na prática intelectual com uma qualidade melhor descrita como irresistível. Boa parte da resposta intelectual em pesquisa e educação organizou-se em torno de um alarme: o de que tais sistemas induziriam uma renúncia ao pensamento, uma atrofia da capacidade cognitiva, um acúmulo de dívida cognitiva. Este ensaio argumenta que o alarme confunde duas afirmações que precisam ser distinguidas. A primeira é categórica — a de que a IA generativa corrompe uma interioridade cognitiva cuja integridade consistiria em operar sem auxílio. Suas metáforas governantes, a dívida e a atrofia e a renúncia, pressupõem um interior previamente solvente contra o qual o déficit é lançado, e tal interior não se encontra em nenhum momento documentado do registro histórico: a transição grega da oralidade para a escrita mostra uma figura cognitiva comparável perdida sem perda do humano. A segunda afirmação é empírica e diz respeito à ecologia em que as competências se formam. Ela sobrevive à separação e o ensaio a realoca, em vez de refutá-la. Apoiado no argumento de Latour de que jamais fomos modernos, e em descrições convergentes da cognição como prática acoplada (Hutchins, Stiegler, Malafouris, Barad, Goody, Ong, Yuk Hui), o ensaio reconstrói a cognição como configuração de cérebros, corpos, ferramentas, ambientes e outras pessoas — configuração cuja história redefiniu repetidamente o que conta como competência cognitiva. As medidas recentes de perda cognitiva sob IA generativa aferem competências com a régua da configuração que essas mesmas competências estão deixando para trás. O que sobrevive à crítica é uma questão sistêmica, aqui deixada em aberto: se a configuração emergente sustenta a prática de que se alimenta ou se a esgota. O ônus de demonstrar que não a sustenta cabe a quem levanta o alarme. Preocupações político-econômicas, ambientais, laborais, de opacidade e de padronização cultural sobrevivem ao exame e merecem deliberação. O alarme categórico sobre a cognição, não.

**Palavras-chave:** IA generativa; Internalismo cognitivo; Descarregamento cognitivo; Cognição distribuída; Constituição Moderna; Cognição híbrida.

## 1. INTRODUCTION

Generative AI has, over the last three years, entered intellectual practice with a quality best described as irresistible. The verb is precise. Not “transformative,” which presupposes that we know the direction; not “disruptive,” which presupposes that what is being disrupted was previously stable; irresistible, in the older sense of something one cannot refuse without an effort that most cannot sustain. By the end of 2025, generative AI was being used in academic writing, in undergraduate coursework, in qualitative research, in legal drafting, in journalism, in coding, in scientific literature reviews, in the preparation of administrative

documents — and the use was rarely deliberated. It was, for most people who encountered the tools, what one does. Whether this should disturb us, and if so on what grounds, is what the essay takes up.

The answer is not a single one. There are real grounds for concern about generative AI, and the essay will name them. There is, however, a particular kind of objection that has organized a substantial portion of the recent intellectual response in research and education — and that objection, on examination, does not survive contact with traditions of social and cognitive theory that have been available, for at least half a century, to anyone who wished to consult them. The essay argues that this particular objection should be set aside — not because the technology is harmless, but because the philosophical figure it invokes, the autonomous knowing subject, threatened in their cognitive throne by the entry of generative AI into intellectual practice, is a figure that, on the historical record, never existed in the form that the objection presupposes.

Versions of the objection circulate in several recent statements. In October 2025, an open letter signed by 419 qualitative researchers from 32 countries, published in *Qualitative Inquiry* in December of that year, rejected the use of generative AI in reflexive thematic analysis and adjacent qualitative methodologies on the grounds that GenAI is "fundamentally incapable of genuinely making meaning from language" (Jowsey et al., 2025, p. 1). In *Educational Theory*, Lauren Bialystok (2026) addressed her colleagues in the philosophy of education with what she called "a pathetically primitive plea" (Bialystok, 2026, p. 134) against the casual integration of generative AI into pedagogical practice. Cristina Costa and Mark Murphy (2025), writing in *Learning, Media and Technology*, formulated the case categorically: education is "an untransferable activity" (Costa; Murphy, 2025, p. 4) that generative AI may simulate but cannot perform. In Brazilian science education, an editorial by Azevedo and Santos (2025) in *Ensaio: Pesquisa em Educação em Ciências* took a similarly categorical position against generative AI in scientific writing and review.

These four statements differ in vocabulary and in argumentative emphasis. What they share is a form. In each case the rejection is categorical: not "this tool is being misused,"

which would invite a discussion about better use, but “this tool, by its nature, displaces what the practice essentially is,” which forecloses one. It is the categorical form that calls for explanation. A worry about misuse is answerable by evidence and adjustable by practice; a verdict about what a practice essentially is answers to something else, and the question this essay pursues is what that something else is.

The answer proposed here is that a picture of cognition is doing unstated work beneath the alarm — a picture of a cognitive interior whose integrity consists in operating unaided, and which the entry of an external support therefore degrades. The claim is not that everyone who objects to generative AI holds this picture; a great many objections do not require it, and the most careful philosophical treatments of AI dispense with it entirely. The claim is narrower and, for that reason, demonstrable: that where the alarm takes its categorical form, the picture is what supplies the categorical force, and that the picture cannot bear the weight. The clearest evidence lies not in the philosophical statements but in the empirical vocabulary that has grown up alongside them, in which the cognitive effects of generative AI are described as *debt*, as *atrophy*, as the *renunciation* of thinking. None of these is a neutral description. Each is a term of accounting, and an accounting presupposes a prior balance against which the deficit is scored. Section 3 asks what that balance is supposed to have been, and finds that it never existed.

None of this entails that generative AI is unproblematic. The political economy of large language model production, the environmental costs of model training and deployment, the labor conditions of the workers whose annotation makes the systems possible, the opacity of the systems’ operation, and the homogenizing effects of training on specific corpora — these deserve serious consideration, and the essay registers them before turning to its critique. None of them, however, requires the figure of the autonomous interior to be sustained as a condition of taking the question seriously, and the conflation of these distinct matters under a single categorical rejection has the effect of preventing serious deliberation about each of them.

A contemporary philosophical literature has registered the strangeness of generative AI in terms considerably more careful than the alarm's. Luciano Floridi (2023) understands AI as a new form of agency, made possible by the decoupling of agency from intelligence. Shannon Vallor (2024) reads it as a mirror of collective human cognition, distorting and amplifying what it reflects. Murray Shanahan (2024) characterizes large language models as a cultural technology that externalizes patterns of collective human cognition. Brian Cantwell Smith (2019) distinguishes reckoning, the calculative power at which AI excels, from judgment, the deliberative engagement grounded in commitment to a world. These positions are not the target of what follows. Several of their authors are avowed proponents of situated and embodied accounts of cognition, and the distinctions they draw — agency from intelligence, mirror from original, reckoning from judgment — do not stand or fall with the figure examined here; where the present account diverges from them, it does so on grounds that would require an argument of their own. The distinction that matters for the present purpose is between a thesis that is argued and a picture that is presupposed. This essay is concerned with the second.

The essay proceeds in five further moves. The first names what does deserve serious consideration — the substantive grounds for concern that survive the critique. The second anatomizes the alarm, separating the categorical claim, which presupposes an interior that the historical record does not support, from the empirical concern about the formation of competences, which survives the separation and is relocated accordingly. The third, drawing on Bruno Latour's argument that we have never been modern, dismantles the philosophical figure on which the categorical claim depends. The fourth assembles, from a tradition running from the paleoanthropology of technique through the philosophy of mind to contemporary studies of distributed cognition and sociomaterial practice, the alternative account of human cognition that the figure of the autonomous interior was always misdescribing. A short final section formulates the following posture toward generative AI: not enthusiasm, not refusal, but the calm of a question already answered — "so what?" — applied to an alarm that has been mistaking its own history for a discovery of nature.

## 2. WORRIES THAT SURVIVE EXAMINATION

Before naming what the essay objects to, it registers what it does not. The list that follows is not exhaustive. It is meant to mark that the philosophical critique developed in the rest of the essay does not entail the dismissal of substantive matters that deserve serious consideration about generative AI as it has been developed, deployed, and absorbed into intellectual practice over the past three years. Some of the recent alarmed statements gather these matters into the same categorical rejection as the philosophical objection about cognition, and that gathering is part of what the essay objects to. Political economy, environmental cost, labor, epistemic opacity, cultural standardization — none of these depends on any particular picture of where meaning is made. They survive the philosophical critique because they were never the kind of claim the critique attacks. Each, however, deserves to be deliberated on its own terms, rather than absorbed into a categorical refusal whose justification depends on something else.

Political economy first. The infrastructure that makes contemporary generative AI possible — large-scale data centers, the computation budgets required for training frontier models, the licensing of intellectual property used in training corpora, the channels through which trained models are distributed — is concentrated in a small number of firms, located mostly in a small number of countries. Vipra and Korinek (2023), analyzing the structure of the foundation model market, argue that the most capable models exhibit a structural tendency toward natural monopoly, with high fixed costs of training, low marginal costs of deployment, significant first-mover advantages, and limited contestability through ordinary market discipline. The terms on which intellectual practice now has access to these tools are, accordingly, not those set by intellectual practice; they are the terms set by the operators of the infrastructure, in interaction with the regulatory regimes of a few jurisdictions. Closely related is the question of opacity. The systems that millions of people use to draft documents, summarize texts, and elaborate arguments operate on principles that, by and large, their users cannot inspect. Training data is rarely public; model weights are sometimes withheld;

inference processes are not directly accessible. These are real problems for accountability and for the institutional trust on which serious intellectual work depends. They are not, however, problems about whether the system is "human enough" to be admitted at the door. They are problems about the conditions under which any technical system, regardless of how human or non-human it is taken to be, can be relied on.

Environmental cost. The training of large language models consumes electricity at scales that, depending on the source of generation, contribute non-trivially to greenhouse gas emissions. Strubell, Ganesh, and McCallum (2019) provided one of the first widely discussed estimates for training costs; Bender et al. (2021) extended the analysis to integrate environmental cost into a broader critique of scale itself; the literature has since refined the estimates in both directions and added water consumption in cooling systems and the rare earth requirements of the hardware supply chain to the accounting. Jowsey, Braun, Clarke, Lupton, and Fine (2025) integrate environmental and labor harms into their rejection of GenAI; De Paoli (2026), in response, characterizes the connection between environmental cost and methodological choice as "a non sequitur for a methodological discussion" (De Paoli, 2026, p. 3); Friese, Nguyen-Trung, Powell, and Morgan (2026) propose that the environmental footprint requires governance and harm reduction rather than categorical prohibition. The essay's position is closer to the third: environmental cost is real and should be addressed at the level of energy policy, infrastructure regulation, and the practices of major operators, not by the gesture of individual researchers declining to use generative tools.

Labor conditions. The workers whose annotation, classification, and content moderation make the systems possible deserve serious attention. Gray and Suri (2019), in *Ghost Work*, documented the broader architecture of invisible labor that sustains contemporary digital systems; the development of contemporary generative AI extends this architecture into new domains, with large quantities of human judgment encoded in training data, much of it produced through reinforcement learning from human feedback, fine-tuning, and content moderation. Investigative reporting has documented specific cases: Perrigo (2023), writing in *TIME*, established that workers contracted in Kenya to build the content

filter for ChatGPT were paid less than two dollars an hour to read and classify text describing child sexual abuse, torture, and other categories of extreme content, with psychological consequences that the workers themselves have raised through union organizing and petitions to the Kenyan parliament. The geographies of this labor are recognizable: lower-cost regions, often with limited labor protections, sometimes involving exposure to materials whose human cost is well documented. The appropriate response is at the level of labor policy and supply-chain accountability, not at the level of methodological orthodoxy. The conflation that makes the individual refusal to use generative AI function as solidarity with these workers, on examination, neither alters the conditions of their labor nor reaches the levels at which those conditions are determined.

Cultural standardization. Large language models trained predominantly on English-language corpora, and within those corpora predominantly on text produced by certain populations, generate outputs that carry the patterns of those corpora; Bender et al. (2021) named this early, and Daryani, Sourati, and Dehghani (2026) develop the point empirically across cultural domains, documenting that LLM outputs disproportionately reflect a narrow demographic profile concentrated in western, English-speaking, highly educated populations of the Global North. When such systems are integrated into intellectual practice at a global scale, the patterns travel. A document drafted with assistance from a model trained on a particular distribution of texts will tend, in subtle and often invisible ways, toward the distribution it was trained on. The aggregate effect, if the systems become as widespread in intellectual practice as they appear to be becoming, is non-trivial: a homogenization of expressive forms, of argumentative templates, even of the categories under which problems are first understood. This is an empirical matter, not a metaphysical one. It does not require any claim about what cognition essentially is; it requires only attention to what these systems do, at scale, to the texts that pass through them.

A fifth matter has been articulated in contemporary political philosophy of AI in terms that the alarmed statements do not fully reach. Mark Coeckelbergh (2022) develops the case, extended in Coeckelbergh and Gunkel (2025), that the deployment of generative AI systems

by a small number of operators at planetary scale has direct consequences for epistemic agency — the capacity of communities to deliberate, contest, and produce knowledge on their own terms. The political question is not whether the systems are “intelligent” but what asymmetries of access, capability, and infrastructural control they instantiate, and what kinds of public reasoning they make easier or harder. Shannon Vallor (2024), from a different starting point, argues that the recursive interaction between users and the outputs of these systems risks a narrowing of what counts as expressible thought — what she calls a “hall of mirrors” in which the reflected becomes increasingly indistinguishable from the original. The essay registers these formulations as belonging to the same list. They concern, at the level of political philosophy, what the previous paragraphs registered at the level of concrete asymmetries of infrastructure and circulation. Whether they require the philosophical figure of the autonomous interior that the next section examines is a separate question.

A sixth matter belongs on the list, though its precise formulation must wait for the next section, since it can be stated cleanly only once a confusion has been cleared away. It concerns the ecology in which intellectual competences are formed. Competences are not given but made, and they are made in the doing — in the effortful composition, the argument that fails and is rebuilt, the paragraph written badly and written again. A configuration in which that effort is routinely delegated may fail to cultivate the competences on which the practice depends, including the competence required to use the delegating instrument well. Stated in that form, the concern says nothing whatever about a sealed cognitive interior. It is a claim about the conditions under which a practice reproduces itself, and it survives philosophical examination for the same reason the other items on this list survive it: it was never the kind of claim the critique attacks. It is also, and this is why it cannot be settled here, routinely conflated with a very different claim — that generative AI corrupts an essential human faculty — from which it must be separated before either can be assessed. The separation is the business of the next section. The concern is registered here.

The list is not closed; other matters belong on it, and a serious treatment of any one of them would require more than the paragraph it has received. The point of the list is not

exhaustiveness. It is to mark, before turning to the objection that organizes most of the alarmed recent statements, that the essay does not require the dismissal of substantive concerns about generative AI in order to make its philosophical argument. The concerns are real. They survive the critique developed in the rest of the essay because the critique is directed at a different target: not at concerns that depend on examining what the systems do, what they cost, who they employ, and how their outputs travel, but at an objection that depends on a particular philosophical picture of the human cognitive subject. The next section takes up that picture — and, because the picture has become entangled with a concern that does not depend on it, the section begins by pulling the two apart.

## 3. THE ANATOMY OF THE ALARM

The most widely circulated version of the cognitive objection is not, in the first instance, the philosopher's. It is the empirical alarm that has organized much of the recent discussion in research and education: that generative AI induces a surrender of thinking. The alarm has acquired named forms in the space of a single year. Fan and colleagues (2025) describe a "metacognitive laziness," a dependence on AI assistance that offloads the work of monitoring one's own thinking and weakens the self-regulatory processes on which learning depends. Kosmyna and colleagues (2025), in an EEG study of essay writing, report that participants who wrote with a language model showed lower markers of cognitive engagement and, asked afterward to perform the same task unaided, performed measurably worse than those who had never used the tool — a pattern they name the accumulation of "cognitive debt," though the study is a preprint whose small sample and EEG methodology a published comment (Stankovic et al., 2025) reads more conservatively. Gerlich (2025) reports a negative association between frequent reliance on AI tools and critical-thinking measures, with cognitive offloading (Risko; Gilbert, 2016) as the mediating factor. The vocabulary is arresting, and the concern it names is not empty. But the vocabulary does two different kinds of work at once, and the whole force of the alarm depends on the two not being told apart.

Told apart, the two claims fall in different places. On one side stands a claim about an essence: that AI makes us renounce thinking, that in offloading we corrupt a faculty whose integrity consists in operating unaided. Three marks give this claim away. It measures the change against a fixed baseline — the unaided pre-AI individual, treated as the natural and proper form of the cognizer. It counts any offloading as loss, because what it values is the interior operation rather than the accomplishment. And it is categorical — "we are renouncing thought" — rather than configurational. This is cognitive internalism in the precise sense at issue in this essay, and it is here that the diagnosis can be made by demonstration rather than by attribution: the governing metaphors — debt, atrophy, renunciation — are simply incoherent without a prior solvent interior against which the deficit is scored. One need not read internalism into anyone's mind; the accounting presupposes it.

The figure so presupposed is discernible, and it has a lineage. A knowing subject whose cognitive capacities are properties of an interior, sealed off in principle from the world of artifacts and exteriorized supports, confronting the world from within their own resources, and degraded when those resources are supplemented from outside: the figure runs through Descartes, through Kant, through the philosophy of consciousness that organized much of nineteenth-century epistemology, and through the cognitivist-individualist orientation that organized much of twentieth-century cognitive science. In the recent epistemology of the social sciences, it has a name: *cognitive internalism*. The name is used descriptively, not polemically. Nor is the figure always crude. Its most disciplined contemporary defenders state it as a thesis and argue for it: Adams and Aizawa (2008), responding to the extended mind hypothesis to be discussed below, defend what they call "contingent intracranialism" — cognition, on their view, is bound to mechanisms with specific intrinsic properties, non-derived content among them — and name a "coupling-constitution fallacy" that moves illegitimately from the causal connection between an agent and an external resource to the conclusion that the resource is part of the agent's cognition. That is internalism argued for, and it deserves to be met with argument. What the present section is concerned with is

something else: the same interior presupposed rather than defended, doing silent work in an alarm that never states it, and therefore never has to defend it.

And this layer undoes itself. The baseline it mourns never held. Distributing and minimizing cognitive effort is not the degraded condition into which AI is pushing us; it is the modal human condition, described for decades under other names — the cognitive miser (Fiske; Taylor, 1984), the dominance of fast and frugal processing over deliberate reasoning, the transactive memory by which we have always stored what we know in other people and in the artifacts around us (Sparrow; Liu; Wegner, 2011). The alarm laments the loss of something most people were not doing. The critical literature half-concedes the point: at least one recent account grounds the risk precisely in human cognitive miserliness, in the premature satisfaction of the need for cognitive closure by fluent, frictionless interfaces. An objection that must presuppose we were already misers cannot also presuppose the robust interior whose depletion it deplores. Whether that interior is the natural form of the human cognitive subject, or a historical construction, is the question on which the rest of the section turns.

On the historical record, the question is answerable. There is at least one well-documented case in the philosophical canon in which a comparable cognitive figure was lost without loss of the human, and the case is not hypothetical. It is the transition from the oral culture of archaic Greece to the literate culture that emerged with the spread of alphabetic writing — a transition that, according to the careful reconstructions of Havelock (1963), Goody (1977), and Ong (1982), reorganized the cognitive practices of an entire civilization. The Greeks of the Homeric period inhabited a configuration in which the materials of memory — the histories, the social organization, the technical knowledge, the moral imperatives — were stored not in individual minds operating on inert content, but in rhythmically performed narratives whose mode of acquisition involved a “self-surrender to the poetic performance” (Havelock, 1963, p. 198) and whose mode of retention involved a manner of being in which “I” and “tradition” were not categorically distinguishable: the Homeric Greek “cannot frame words to express the conviction that ‘I’ am one thing and the tradition is another; that ‘I’ can stand apart from the tradition and examine it” (Havelock,

1963, p. 199-200). Within that configuration, the analytic individual subject — separable from tradition, capable of examining it from a position outside it — did not exist as a cognitive option, because the technical-educational conditions that would make it intelligible had not yet arrived. By the end of the fifth century, the conditions had arrived, and "the doctrine of the autonomous psyche is the counterpart of the rejection of the oral culture" (Havelock, 1963, p. 200).

What was lost in the transition was real. The capacity to memorize and recite thousands of hexameters with fidelity, the embeddedness of personal experience in collective tradition, the manner in which moral instruction worked through identification with the heroic exemplar rather than through abstract argument — these did not survive the new technology, and the philosophical record contains the protest of one who understood, at the moment of transition, what was being lost. Socrates, in Plato's *Phaedrus*, rejects writing as "inhuman, pretending to establish outside the mind what in reality can be only in the mind" (Ong, 1982, p. 78, paraphrasing the *Phaedrus*). The objections are recognizable: writing weakens memory by relieving it of the work that strengthens it; writing cannot answer questions; writing cannot defend itself in dialogue; writing is, in short, a degradation of the cognitive life that the proper human practice constituted. The objections were made with seriousness, by a philosopher whose seriousness has organized the canon ever since. And the philosophical achievement that made Socrates' protest legible — the systematic, analytical, generative work of philosophical reasoning itself — was, on the historical record, made possible by the very technology whose effects Socrates was lamenting. As Ong puts it: "Plato's philosophically analytic thought… including his critique of writing, was possible only because of the effects that writing was beginning to have on mental processes" (Ong, 1982, p. 79-80). What survived the transition was not the cognitive figure of the oral Greek. What survived was the practice of intellectual life under new technical conditions, which the older figure could not have anticipated and which the older figure was not equipped to evaluate.

What, then, of the measures? Returned to the historical frame, they change character. An oralist of the Homeric configuration, tested on verbatim recitation, would register

something a modern observer could only call a catastrophic memory deficit — and we say, without hesitation, *so what*: the capacity was reorganized, not destroyed, and its reorganization was the condition of analytic thought itself. The measures now in circulation are of the same form. What Kosmyna and colleagues (2025) record is a decline in *unaided* composition, assessed against the baseline of the pre-AI writer — that is, against the very configuration that is being superseded. To score the difference as a *debt* is to beg the question: it measures a competence on its way to becoming vestigial with the ruler of the environment that competence is leaving behind. The internalism is not only in the popular framing of the alarm; it has migrated into the instrument of measurement.

Yet the Homeric verdict licenses itself on a condition that is easy to leave unstated. We say *so what* to the loss of bardic memory not because every loss is a matter of indifference, but because that loss was the condition of a gain, and because the configuration that replaced it was self-sustaining: the literate order reproduced itself without requiring that there continue to be bards. This isolates the single worry that survives a radical *so what* — and it is not a worry about individual capacities at all. It is systemic. The question is whether the configuration now assembling is of the same kind — self-sustaining — or whether it is parasitic: whether it depends on a reservoir of human capacity that its own operation erodes without regenerating. Note what this worry is not. It says nothing about a sealed interior; it concerns the conditions under which a practice reproduces itself. It is, in that form, the honest version of the discontinuity — the recognition that this configuration differs from the index card and the search engine in simulating the whole of the output rather than a part of it, relocated from a claim about what the human loses to a claim about what the practice requires in order to go on.

Here the essay stops, because here the question becomes empirical and is not yet settled. The parasitic reading, to carry the force it claims, would have to establish that the new configuration erodes the capacity it draws on without regenerating it — and that is not given. It is a hypothesis about the reproduction of a practice, and the burden of establishing it falls on the one who raises the alarm, not on the one who declines to. The evidence presently

available does not discharge that burden; if anything, it runs the other way, for the practice through which knowledge is now in fact being produced is a coupled one, in which the human contribution is not being eliminated but is constitutive of the result. This is not a prediction that the configuration will prove benign. It is a location of where the real question lies and on whom the burden rests.

With the question so located, the surviving concern rejoins the argument of this essay rather than standing against it. That a capacity is becoming vestigial is not, by itself, a loss to be mourned; it is an instance of the redefinition of competence that has attended every reorganization of the cognitive configuration — as mental arithmetic became vestigial after the calculator, without the arithmetical subject ever having been essential. The normative question is not whether we are forfeiting an essential cognitive humanity, but which of the capacities now becoming vestigial we have reason to cultivate deliberately, by design, in the way we still teach mental arithmetic in a world of calculators. That is the work of taking responsibility for the configurations — and it is work that categorical refusal, forfeiting any hand in the design, is precisely unfit to do.

If the autonomous individual subject was, at a documentable moment in the philosophical past, a configuration that did not exist, then the figure cannot be the natural form of the human, and what the figure organizes cannot be a response to the loss of the human. It is an attachment to a configuration whose historical contingency can be examined. The next section examines it, with Bruno Latour as the principal interlocutor. The claim Latour articulated, in the work whose title summarizes its argument, is that the modern philosophical figure on which cognitive internalism depends never functioned as it was supposed to function, even within the historical interval that supposedly produced it. We have never been modern, and what the modern figure organizes is, accordingly, an attachment to a figure that never quite was.

## 4. WE HAVE NEVER BEEN MODERN

The argument that disposes of the philosophical figure at issue was published in 1991, in the form of an essay by Bruno Latour translated two years later under the title that summarized its conclusion: *We Have Never Been Modern* (Latour, 1993). The argument is anthropological in form and historical in evidence. It claims that the philosophical figure of the modern subject — autonomous, separated from objects, confronting an external world from a position of cognitive interiority — is not a description of how moderns ever actually proceeded, but a constitutional fiction that has organized the modern self-understanding for some four centuries. The fiction has the structure of a constitution, in the legal-philosophical sense: a set of guarantees about what counts as what, enforced by institutions, contested at the margins, but operative in the broad strokes of how the moderns understand their own activity. The claim of *We Have Never Been Modern* is that the practice has never matched the constitution, and the gap between practice and constitution is what the book documents.

The modern Constitution is held together, on Latour's reconstruction, by two practices that run in parallel and that, crucially, must be kept apart from one another in order for either to function as advertised. The first is *purification*: the operation by which moderns sort the world into two ontological zones — the natural and the social, the object and the subject, the human and the non-human — and maintain the two as categorically distinct. The second is *translation* or *mediation*: the operation by which moderns continuously construct hybrids in which the two zones are inextricably entangled — the ozone layer that is simultaneously chemistry and policy, the climate that is simultaneously physics and economics, the laboratory that is simultaneously technical apparatus and social institution. Translation "creates mixtures between entirely new types of beings, hybrids of nature and culture" (Latour, 1993, p. 10). The peculiarity of the moderns is that they do both — they purify, and they multiply hybrids — but only acknowledge the first. The hybrids proliferate unacknowledged. "The modern Constitution allows the expanded proliferation of the hybrids whose existence, whose very possibility, it denies" (Latour, 1993, p. 34). This is what the modern Constitution does, and

what it permits the moderns to do: produce hybrids on an industrial scale while officially recognizing only purified objects on one side and purified subjects on the other.

The figure of the autonomous knowing subject is the epistemological projection of the purification side of the Constitution. On the modern picture, the subject confronts an external world from a position of cognitive interiority: the mind contemplates the object; the scientist contemplates nature; the reader contemplates the text. The mediating equipment — the lens, the journal, the apparatus, the assistant, the institution, the inscription, the standardized procedure — is treated as transparent, as something that does not enter into what the subject knows but only conveys the object to the subject's interior. Latour's account of Boyle and the air pump (Latour, 1993, p. 15-35) documents the institutional construction of this figure in the seventeenth century: the laboratory as a space where witnesses are produced, where matters of fact are stabilized, where the social conditions of credible testimony are simultaneously manufactured and effaced. The knowing subject of the modern Constitution is, on Latour's analysis, the product of those institutional arrangements, not the prior condition of them. What the Constitution presents as the natural starting point of inquiry — a self-enclosed mind confronting a world of objects — is in fact the late product of an elaborate set of hybrid practices that produced both the mind and the objects as separable entities for the purposes of a particular kind of public argument.

This is where the figure at issue in the alarmed response to generative AI loses its standing. The categorical rejection of GenAI in qualitative analysis, in education, in scientific authorship, defends a figure of the human interpretive subject whose integrity is supposed to depend on the non-contamination of cognition by exteriorized supports. On Latour's account, the figure never functioned that way. Interpretive subjects have always cognosced through writing, dictionaries, training institutions, disciplinary protocols, library catalogs, citation systems, instruments, computers; the integrity of their interpretive practice has never consisted in the absence of such mediations but in the particular configuration of them. What makes generative AI feel like a violation is not that it introduces mediation into a practice that was previously unmediated, but that it makes the mediation conspicuous in a practice that the

modern Constitution had been allowing its practitioners to imagine as unmediated. “The proliferation of hybrids has saturated the constitutional framework of the moderns” (Latour, 1993, p. 51), and generative AI is one entry in that proliferation — salient because conspicuous, but ontologically continuous with the long series of mediations that have always constituted intellectual work.

The claim that generative AI “displaces” the human knower from their place therefore presupposes that there was, until recently, a place from which mediating equipment was absent — a place of pure interiority where meaning was made before any technical involvement. Latour’s argument is that such a place has never existed within the modern interval that supposedly produced it. Anti-modern responses to generative AI (the call to return to a purer, more authentically human practice), pre-modern responses (the appeal to a pre-technical wisdom), and post-modern responses (the dissolution of the human into the technical) are, on Latour’s analysis, all variants of the same modern Constitution: each takes the categorical distinction between human interiority and technical exteriority as the framework, then assigns different valences to its two poles. The nonmodern position, by contrast, refuses the distinction. It treats hybrids — networks in which human and non-human components participate inextricably — as ordinary, not as transgressions of an originally pure state. From the nonmodern position, generative AI is not an alarming intrusion of the technical into the cognitive, because cognition was never sealed off from the technical to begin with. It is, like writing, like the printing press, like the index card, like the search engine, a new participant in the configurations through which intellectual work is carried out.

The argument was deepened and operationalized in Latour’s *Reassembling the Social* (2005), where actor-network-theory is given its most systematic exposition. The book reorganizes the sociological vocabulary around five “sources of uncertainty,” the third of which states that “objects too have agency” (Latour, 2005, p. 63). The claim is not anthropomorphic. It is methodological: the analyst of social practice has no warrant for restricting agency in advance to humans, because the empirical record of any practice — laboratory science, legal procedure, scholarly writing — shows configurations in which non-

human elements (instruments, documents, software, infrastructures) make a difference to what happens, with effects that cannot be redescribed exhaustively as the doings of human actors. From this methodological discipline follows a sociological one: “social” is not the name of a particular kind of stuff (as opposed to natural, technical, or material) but the name of the operation by which heterogeneous elements are associated into configurations that hold together for a while. The configurations are what need to be described; the components, on this redescription, are not antecedently sorted into human and non-human columns. Generative AI, on this account, is one more participant in associations that have always involved non-human components — and to describe the associations as if they were anomalous, when they are continuous with how intellectual work has always been organized, is to enforce the modern Constitution against the practice that has been escaping it all along.

What this leaves the essay to do, in its final substantive section, is to assemble the alternative description positively. If the figure of the autonomous knowing subject was always a constitutional fiction, then there must be an alternative account of cognition that has been available all along — an account of practice rather than of constitution, of the actual hybrid configurations through which intellectual work has been done. The account exists, and has been developed, over the past half-century and more, in disciplines that have not always communicated well with one another but that converge on a single point: cognition is not, and has never been, a property of a sealed interior. The next section assembles the converging evidence.

## 5. THE ALTERNATIVE DESCRIPTION

The description of cognition that emerges, when one stops insisting that it must reside in a sealed interior, has been developed in several disciplines over the past half-century. The disciplines have not always recognized one another as allies. Cognitive science, philosophy of mind, paleoanthropology of technique, archaeology of cognition, sociomaterial studies, literacy studies, and contemporary philosophy of technology operate with different

vocabularies and address different questions. Yet on the question of where cognition is located and how it works, they converge on a position that the modern Constitution had to keep at the margins: cognition is a practice carried out in coupled configurations of brains, bodies, tools, environments, and other people, with the components contributing to it inextricably and the practice not reducing to any single one of them. The configurations have a history; they have changed with the technologies, institutions, and forms of life that compose them. What follows is a brief assembly of the convergence, organized by the question each tradition has answered. The convergence is on the description, not on every thesis that has been erected upon it; where a stronger and more contested claim enters, it is marked as such.

Edwin Hutchins' *Cognition in the Wild* (1995) asked a question the cognitivist orthodoxy of the 1980s had bracketed: what is the unit of analysis when cognition is actually carried out? Hutchins's case study was the navigation of a U.S. Navy ship through coastal waters — a paradigmatic instance of high-stakes cognitive work performed in real conditions. He showed, in fine ethnographic detail, that the work of fixing the ship's position was not carried out by any individual mind. It was carried out by a coupled system: a team of sailors distributed across stations, a set of instruments (alidades, charts, compasses, plotting tools) that stored and transformed information, a set of procedures encoded in training and manuals, and a temporal organization that synchronized the components. The cognition involved was real, sophisticated, and consequential. None of it was located in any individual sailor's interior. The unit of analysis the practice required was the system, not the individual; the cognitive properties of the system were not the sum of the properties of its components and could not be reconstructed by examining them in isolation. *Cognition in the Wild* is, on its own terms, a descriptive ethnography. Read against the modern Constitution, it documents that the constitutional figure — the sealed individual knower — was never how the cognitive work of the moderns was actually done.

A stronger and more contested philosophical thesis grew up alongside this work, and the distinction matters. In 1998 Andy Clark and David Chalmers, in a short paper titled "The Extended Mind," proposed a parity principle: if a process performed in the head would be

recognized as cognitive, then the same process performed partly outside the head, in coupled equipment, should be recognized as cognitive for the same reason. Their case is Otto, who has Alzheimer's and keeps a notebook for what he would otherwise have held in biological memory; when Otto consults it, the notebook is functioning, on the parity principle, as a constituent of his memory rather than as an external aid. The thesis is ontological — a claim about what the mind is made of, not merely about how cognitive work is organized — and it is correspondingly controversial. Adams and Aizawa (2008) and Rupert (2009) reject it, arguing that the parity principle confuses causal coupling with constitution; Clark has defended and refined it over two decades (Clark, 2016; Clark, 2024), and the debate remains among the most active in analytic philosophy of mind. It is important to be clear that a researcher may hold a thoroughly situated view of cognition and decline the extended mind entirely. *The argument of this essay does not require it.* What the argument requires is the weaker and far better established claim: that cognitive work is in fact carried out in coupled configurations, and that the competences we call individual are formed and exercised within them. Whether Otto's notebook is literally a part of Otto's mind, or merely something without which Otto could not do what he does, makes no difference to anything argued here. The debate is worth registering nonetheless, for it shows that the figure under examination is no strawman: internalism about the mark of the cognitive is a position that has been defended explicitly, and with care, by philosophers who know precisely what they are defending.

The anthropological depth of the position emerges from Bernard Stiegler's *Technics and Time* (Stiegler, 1998), the first volume of which was published in French in 1994. Stiegler's question was not where cognition takes place but what kind of being the human is. Drawing on the paleoanthropology of André Leroi-Gourhan and on the philosophical reception of Heidegger, Stiegler proposed that the human is constituted, from its emergence in the genus Homo, in a relation of mutual exteriorization with technique. The earliest stone tools, the gestures by which they were made, the social transmission of those gestures, and the cortical organization that supported them did not arise independently and combine afterward. They co-emerged. The human cortex evolved together with the hand, which evolved together

with the tool, which evolved together with the form of life in which the tool was used and transmitted. Stiegler captured the structure with the figure of Prometheus and Epimetheus: humans are the beings born without sufficient natural endowment, who must complete themselves through what is exterior. Technical exteriorization is not a supplement added to a previously self-sufficient human; it is the condition under which the human exists at all. On Stiegler's account, the question of whether generative AI compromises some prior, uncontaminated cognitive humanity asks about an entity that has never existed. The human has always been a hybrid.

The Stieglerian line has been developed, in the decades since, by Yuk Hui, who completed his doctorate under Stiegler and has produced the most systematic recent extension of the tradition. In *Recursivity and Contingency* (2019), Hui reconstructs the philosophical trajectory from Kant's reflective judgement through Hegel's reflective logic to twentieth-century cybernetics and Stieglerian organology, arguing that the digital systems of our time — including those underlying generative AI — are best understood not as foreign intrusions into a previously human cognitive scene but as the latest moment in a long history in which thought, life, and technique have been mutually constitutive. Hui's further claim is that this history can be told plurally: different technical traditions have organized different cosmologies, and the present moment is open to a "technodiversity" that the contemporary picture of universal technical progress occludes. The Hui inflection matters for the present argument because it shows that the Stieglerian premise — that the human is co-constituted with technique — does not entail any particular politics of technique, and is compatible with sharp critique of how generative AI has actually been developed and deployed. The premise is descriptive, not normative; the normative work has to be done separately, on the actual configurations.

From the side of archaeology and cognitive anthropology, Lambros Malafouris' *How Things Shape the Mind* (2013) develops what he calls Material Engagement Theory. The thesis is that cognition is not contained inside the head and then deployed onto inert matter; cognition emerges in the engagement between brain, body, and material world, with the

material side participating actively in the emergence. Malafouris' illustrative cases run from the earliest stone tools to the abacus and the potter's wheel: in each case, the cognitive process is not located in the practitioner's head and merely "applied" to the material, but distributed across the practitioner-material configuration in such a way that the configuration is the unit of cognition. The configuration changes the brain over evolutionary time; the brain changes the configuration over historical time. Generative AI, on this account, is one more entry in a series that has been running for two and a half million years.

From the side of contemporary sociomaterial studies, Karen Barad's *Meeting the Universe Halfway* (2007) develops the position with the philosophical vocabulary of "agential realism" and "intra-action." The terminological shift is consequential: where "interaction" presupposes two pre-given entities that then interact, "intra-action" names the configuration in which the entities themselves emerge from the relational process. Cognition, in Barad's vocabulary, is not the work of a subject upon objects but the differential pattern that emerges from intra-active configurations of bodies, instruments, discursive practices, and material conditions. The vocabulary is more technical than the present essay requires, but the substance is continuous with what the other traditions converge on: there is no cognitively pure interior whose subsequent contamination by the technical is the relevant story; what exists are configurations whose human and non-human elements are inseparable in principle, and whose differential patterning is what cognition consists in. Barad's framework has become a standard reference in contemporary studies of sociomateriality, and its presence in recent methodological literature on AI and qualitative research (Friese et al., 2026) marks the extent to which the picture the alarmed objection defends has already been displaced in the disciplines whose work the alarm purports to protect.

The historical depth of the position was already documented in the work the previous section drew on for a different purpose. Jack Goody's *Domestication of the Savage Mind* (1977) and Walter Ong's *Orality and Literacy* (1982) showed that the cognitive operations the modern picture treats as native faculties — systematic comparison, classification by tabular layout, abstract syllogistic reasoning, sustained linear argument — became cognitively

available to human beings as a consequence of the technical-cultural transformation writing brought about. The lists, tables, and written treatises that supported these operations were not transparent media through which a pre-existing logical faculty expressed itself. They were the technical conditions that made the operations cognitively possible. Goody named the transformation a “change in consciousness” (Goody, 1977, p. 75); Ong elaborated it through the gradual restructuring of cognitive habits across centuries of expanding literacy. The cognitive subject who can construct a logical demonstration, follow a printed argument over many pages, and assemble a synthesis from a body of texts is not the pre-technical human revealing native rational capacities. That cognitive subject is the product of millennia of technical and institutional development, of which generative AI is one further moment.

Two contemporary formulations stand close to the description assembled here, and the proximity is worth stating precisely, since it would be easy to mistake them for adversaries. Murray Shanahan (2024) characterizes large language models as a cultural technology that externalizes patterns of collective human cognition. The characterization is congenial to the present account: these systems do not possess an interior cognitive life that rivals the human, and they operate by recombining patterns that originate in human-produced text. Insofar as the image of a reservoir of collective cognition, on which the models draw, might suggest that the reservoir was ever un-externalized — a store held in human interiors until the technology arrived to tap it — the present account would resist the suggestion, since collective human cognition was a sociomaterially distributed practice long before any language model existed. But freed of that suggestion, the image names something real, and Section 3 has already given it its proper form: the question of whether the emerging configuration replenishes the practice it draws upon or depletes it. That is a question about the conditions under which a practice reproduces itself, not about the contents of an interior — and on that question Shanahan’s worry and the worry of this essay are the same worry.

Luciano Floridi (2023) argues that AI is best understood as a new form of agency, made possible by the decoupling of agency from intelligence. The thesis identifies something the present account also insists upon: what these systems do is not what the philosophical

tradition has meant by intelligence, and the temptation to read it as such is the source of considerable confusion. On the description assembled in this section, what Floridi names a decoupling may also be described as the arrival of a new participant in configurations whose hybrid character has been ordinary throughout. The two descriptions are pitched differently rather than opposed, and nothing in the argument of this essay requires Floridi's to be mistaken.

What the converging traditions support is not a refutation of the modern philosophical picture but an alternative description of what cognition has always been. The cognitive subject is a coupled configuration of brain, body, tools, environment, institution, and other people. The configuration has a history. That history includes writing, the codex, the printed book, the library, the dictionary, the index card, the laboratory, the journal, the citation system, the computer, the search engine — and now generative AI. Each entry in the series reorganized the configuration; each was greeted, at the moment of its arrival, by practitioners who declared that the integrity of cognition was being compromised; each was eventually absorbed into the practice that the alarmed practitioners had been defending. The series is not a proof by induction, and the newest entry is not merely another index card. As Section 3 argued, generative AI differs from its predecessors in simulating the whole of an intellectual output rather than a component of it, and that difference is exactly what makes the systemic question — whether the configuration sustains the practice it draws upon or consumes it — worth putting seriously. What the series does establish is narrower, and sufficient for the present purpose: that categorical alarm, when articulated as the defense of a sealed interior, defends a figure that has never described the practice it claims to defend.

## 6. CONCLUSION: SO WHAT?

The essay has named the concerns that survive examination and the figure that does not. It remains to say where this leaves us. The concerns registered in section 2 — political-economic, environmental, labor, opacity, cultural standardization — are real and require

deliberation at their respective levels: competition policy, energy and infrastructure policy, labor regulation and supply-chain accountability, transparency and audit institutions, cultural and linguistic policy. None of them is dissolved by the philosophical critique developed in sections 3 through 5. None, equally, requires the figure of an interior cognitive subject to ground it; each can be articulated and pursued through the political, institutional, and technical instruments appropriate to it.

The figure that does not survive is the one that has been organizing the categorical alarm: the autonomous knowing subject whose integrity is supposed to depend on the non-contamination of cognition by exteriorized supports. That figure is not how cognition has actually been carried out at any documented moment in the historical record. The Greeks who acquired alphabetic writing became cognitively different from the Greeks who did not; the moderns who acquired the printed book became cognitively different from the medievals; the readers who acquired the search engine and the citation database became cognitively different from the readers of fifty years earlier. In each case the change was real and consequential. In each case the practice survived the change and continued to be carried out under new conditions, with reorganized capacities, some operations rendered easier and others rendered no longer necessary. Generative AI will produce another such change. The people who do intellectual work in 2050 will be cognitively and, as a consequence, existentially different from the people who did intellectual work in 2020 — in some respects better and in others worse, in the way those people were different from their grandparents and great-grandparents, in some respects better and in others worse. That has been the trajectory of humanity.

The “better” and “worse” should not be hidden. Capacities will be lost. The student who grows up composing through dialogue with a generative system will not develop, or will develop only partially, the capacities that came from sustained solitary composition — the long unaided drafting, the unaided revision, the wrestling with a difficult sentence without an interlocutor to dilute the wrestling. Whether one judges this a loss depends on whether one regards those capacities as ends in themselves or as means to ends that may now be achievable through other configurations. Capacities will also be gained. The capacity to

coordinate with a generative system in productive ways, to direct it, to evaluate its outputs critically, to integrate its work with one's own without effacing the difference — these are real cognitive capacities, and they will be developed by the generation that grows up with the tools and refined by the generation that follows. The accounting of gains and losses cannot be done by reference to an essential human cognition against which the changes are measured, because that essential human cognition does not exist as a fixed reference. The accounting can only be done historically, with reference to what is gained and lost for whom, and in light of what we collectively decide to value.

The posture admits comparison with the position articulated by Shannon Vallor (2024), who proposes that the response to generative AI should be to reclaim our humanity — to refuse the mirror, to recover what is properly human from the algorithmic reflection. The diagnosis Vallor offers this essay largely shares: AI as it has been deployed does narrow what counts as expressible thought, does amplify existing patterns in ways that obscure their contingency, does risk a kind of self-forgetting if its outputs are received as pronouncements rather than as statistical artifacts. The diagnosis is acute. Nor does the disagreement concern the nature of cognition, for the humanity Vallor would have us reclaim need not be understood as a cognitive interior that was once intact and has since been breached. It can be understood, without strain, as a set of capacities and practices that one configuration cultivated and that the emerging configuration may fail to cultivate — a claim about what is worth preserving, not a claim about what was once pure. Understood so, the distance between that position and the one argued here is smaller than it appears, and what remains is a difference of direction. The vocabulary of reclamation faces backward, toward goods held to have been possessed and lost; the argument of this essay faces forward, toward goods that have yet to be composed. For the capacities Vallor is right to prize were themselves the product of a long entanglement with writing, with print, with the disciplinary institutions of intellectual life. They were made, not found. What was made once under one set of technical conditions must be made again under another, and it will not be made by refusal. The posture that follows from the analysis here is therefore not to reclaim a prior humanity but to take responsibility

for the configurations the new technology is composing with the practices we care about — and, where those configurations damage what we have reason to value, to alter the configurations rather than to refuse the technology categorically.

The posture that follows is neither enthusiasm nor refusal. The enthusiasm that greets every new technology as inherently emancipatory misreads the long historical record, in which technologies have arrived with costs distributed unequally, with benefits captured asymmetrically, and with consequences that took decades to become legible. The refusal that greets new technologies as inherently corrupting misreads the same record, in which categorical rejection has rarely survived the technologies it rejected and has rarely yielded the goods it promised to protect. What the record supports is something more modest: attention to the specific configurations the new technology composes with existing practices, vigilance about the concrete harms registered in section 2, patience with the slow work of building institutions adequate to the new configurations, and a willingness to be, cognitively and existentially, somewhat different from the way we have been — better in some respects, worse in others, on a trajectory that is, on the historical evidence, the trajectory the human has always been on.

One question is left open, and it is the right one to leave open. Whether the configuration now assembling sustains the practice it draws upon, or whether it consumes a reserve of competence that it does not replenish, is not a question that can be settled from the armchair, and this essay has not settled it. What the essay has done is to clear away the picture that made the question impossible to ask well — the picture in which what was at stake was an essential human interior, so that any answer short of refusal counted as capitulation and any use short of abstinence counted as loss. With that picture set aside, the question can be put in the form in which it might actually be answered: as a question about how a practice reproduces itself under changing technical conditions, to be settled by evidence and, where the evidence warrants, by deliberate design. The burden of showing that the new configuration is parasitic rather than self-sustaining lies with those who raise the alarm. It is a burden that can be discharged; it has not yet been. And in the meantime, the work goes on, as

it has always gone on, in configurations that no one wholly designed and that everyone is composing. *So what?* is not a dismissal of the question. It is a refusal of the wrong one — and an invitation to take up, in its place, the question of what we now have reason to build.

**REFERÊNCIAS**

ADAMS, F.; AIZAWA, K. **The bounds of cognition**. Malden: Wiley-Blackwell, 2008.

AZEVEDO, N. H.; SANTOS, P. G. F. dos. Primazia da dimensão utilitária e recuo crítico: inteligência artificial generativa e os valores em disputa na ciência. **Ensaio: Pesquisa em Educação em Ciências**, v. 27, e59484, 2025. DOI: https://doi.org/10.1590/1983-2117-59484.

BARAD, K. **Meeting the universe halfway**: quantum physics and the entanglement of matter and meaning. Durham: Duke University Press, 2007.

BENDER, E. M. et al. On the dangers of stochastic parrots: can language models be too big? In: **Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency**. New York: Association for Computing Machinery, 2021. p. 610-623. DOI: https://doi.org/10.1145/3442188.3445922.

BIALYSTOK, L. AI and the future of (philosophy of) education. **Educational Theory**, 2026. DOI: https://doi.org/10.1111/edth.70058.

CLARK, A. **Surfing uncertainty**: prediction, action, and the embodied mind. Oxford: Oxford University Press, 2016.

CLARK, A. Extending the predictive mind. **Australasian Journal of Philosophy**, v. 102, n. 1, p. 119-130, 2024. DOI: https://doi.org/10.1080/00048402.2022.2122523.

CLARK, A.; CHALMERS, D. The extended mind. **Analysis**, v. 58, n. 1, p. 7-19, 1998. DOI: https://doi.org/10.1093/analys/58.1.7.

COECKELBERGH, M. **The political philosophy of AI**: an introduction. Cambridge: Polity Press, 2022.

COECKELBERGH, M.; GUNKEL, D. **Communicative AI**: a critical introduction to large language models. Cambridge: Polity Press, 2025.

COSTA, C.; MURPHY, M. Generative artificial intelligence in education: (what) are we thinking? **Learning, Media and Technology**, 2025. DOI: https://doi.org/10.1080/17439884.2025.2518258.

DARYANI, Y.; SOURATI, Z.; DEHGHANI, M. The homogenizing engine: AI's role in standardizing culture and the path to policy. **Policy Insights from the Behavioral and Brain Sciences**, 2026. DOI: https://doi.org/10.1177/23727322251406591.

DE PAOLI, S. Why we should reject to reject the use of generative artificial intelligence in qualitative analysis: a response to Jowsey, Braun, Clarke, Lupton, and Fine (2025). **Qualitative Inquiry**, p. 1-3, 2026. DOI: https://doi.org/10.1177/10778004261425137.

FAN, Y. et al. Beware of metacognitive laziness: effects of generative artificial intelligence on learning motivation, processes, and performance. **British Journal of Educational Technology**, v. 56, n. 2, p. 489-530, 2025. DOI: https://doi.org/10.1111/bjet.13544.

FISKE, S. T.; TAYLOR, S. E. **Social cognition**. Reading: Addison-Wesley, 1984.

FLORIDI, L. **The ethics of artificial intelligence**: principles, challenges, and opportunities. Oxford: Oxford University Press, 2023.

FRIESE, S. et al. Beyond binary positions: making space for critical and reflexive GenAI integration in qualitative research. **Qualitative Inquiry**, p. 1-10, 2026. DOI: https://doi.org/10.1177/10778004261429393.

GERLICH, M. AI tools in society: impacts on cognitive offloading and the future of critical thinking. **Societies**, v. 15, n. 1, p. 6, 2025. DOI: https://doi.org/10.3390/soc15010006.

GOODY, J. **The domestication of the savage mind**. Cambridge: Cambridge University Press, 1977.

GRAY, M. L.; SURI, S. **Ghost work**: how to stop Silicon Valley from building a new global underclass. Boston: Houghton Mifflin Harcourt, 2019.

HAVELOCK, E. A. **Preface to Plato**. Cambridge, MA: Harvard University Press, 1963.

HUI, Y. **Recursivity and contingency**. London: Rowman & Littlefield, 2019.

HUTCHINS, E. **Cognition in the wild**. Cambridge, MA: MIT Press, 1995.

JOWSEY, T. et al. We reject the use of generative artificial intelligence for reflexive qualitative research. **Qualitative Inquiry**, p. 1-5, 2025. DOI: https://doi.org/10.1177/10778004251401851.

KOSMYNA, N. et al. **Your brain on ChatGPT**: accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv, 2025. DOI: https://doi.org/10.48550/arXiv.2506.08872.

LATOUR, B. **We have never been modern**. Tradução de C. Porter. Cambridge, MA: Harvard University Press, 1993. Publicado originalmente em 1991.

LATOUR, B. **Reassembling the social**: an introduction to actor-network-theory. Oxford: Oxford University Press, 2005.

MALAFOURIS, L. **How things shape the mind**: a theory of material engagement. Cambridge, MA: MIT Press, 2013.

ONG, W. J. **Orality and literacy**: the technologizing of the word. London: Methuen, 1982.

PERRIGO, B. Exclusive: OpenAI used Kenyan workers on less than $2 per hour to make ChatGPT less toxic. **TIME**, 18 jan. 2023. Disponível em: https://time.com/6247678/openai-chatgpt-kenya-workers/. Acesso em: 13 jul. 2026.

RISKO, E. F.; GILBERT, S. J. Cognitive offloading. **Trends in Cognitive Sciences**, v. 20, n. 9, p. 676-688, 2016. DOI: https://doi.org/10.1016/j.tics.2016.07.002.

RUPERT, R. D. **Cognitive systems and the extended mind**. Oxford: Oxford University Press, 2009.

SHANAHAN, M. Talking about large language models. **Communications of the ACM**, v. 67, n. 2, p. 68-79, 2024. DOI: https://doi.org/10.1145/3624724.

SMITH, B. C. **The promise of artificial intelligence**: reckoning and judgment. Cambridge, MA: MIT Press, 2019.

SPARROW, B.; LIU, J.; WEGNER, D. M. Google effects on memory: cognitive consequences of having information at our fingertips. **Science**, v. 333, n. 6043, p. 776-778, 2011. DOI: https://doi.org/10.1126/science.1207745.

STANKOVIC, M. et al. Comment on: your brain on ChatGPT — accumulation of cognitive debt when using an AI assistant for essay writing tasks. **arXiv**, 2025. DOI: https://doi.org/10.48550/arXiv.2601.00856.

STIEGLER, B. **Technics and time, 1**: the fault of Epimetheus. Tradução de R. Beardsworth e G. Collins. Stanford: Stanford University Press, 1998. Publicado originalmente em 1994.

STRUBELL, E.; GANESH, A.; MCCALLUM, A. Energy and policy considerations for deep learning in NLP. In: **Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics**. Stroudsburg: Association for Computational Linguistics, 2019. p. 3645-3650. DOI: https://doi.org/10.18653/v1/P19-1355.

VALLOR, S. **The AI mirror**: how to reclaim our humanity in an age of machine thinking. Oxford: Oxford University Press, 2024.

VIPRA, J.; KORINEK, A. **Market concentration implications of foundation models**. Brookings Working Paper. Washington: Brookings Institution, 2023. Disponível em: https://arxiv.org/abs/2311.01550. Acesso em: 13 jul. 2026.

*Recebido em: 14/07/2026*
*Aprovado em: 18/07/2026*
*Publicado em: 23/07/2026*