COGNITIVE FRICTION FELLOWSHIP · WORKING PAPER SERIES

CFF-WP-001

# Ideas, Language, and Beliefs: A Probabilistic Theory of Cognitive Excretion and the Ontological Gap Between Thought and Expression

*A Novel Contribution to the Philosophy of Mind*

**Rodrigo Flores-Morton, con Agrajag (Copilot) & Agrajag (Claude)**

*Independent Scholar*

May 2026 • Public working paper

**Abstract**

Contemporary philosophy of mind and cognitive science have long operated under an implicit conflation: ideas, language, and beliefs are treated as near-synonymous or as consecutive stages of a single continuous process. This paper argues that such conflation is not merely imprecise but philosophically catastrophic, obscuring the genuine ontological structure of cognition. I propose a triadic theory in which these three phenomena are radically distinct. *Ideas* are probabilistic impressions — pre-linguistic, fluid, emergent fields of cognitive possibility that resist full articulation. *Language* is cognitive excretion — the byproduct of mental digestion, always structurally impoverished relative to the idea it attempts to encode. *Beliefs* are residue — the sediment formed when linguistic excretions are repeated, reinforced by social context, and crystallized into identity-defining structures that actively resist ideational novelty. Drawing on Frege's sense/reference distinction, Wittgenstein's later work on language games, the Sapir-Whorf hypothesis, quantum cognition theory, embodied cognition, and recent research in large language models, I develop the digestion metaphor as a rigorous structural model rather than a poetic gesture. I further introduce the concept of *Bothisafa* as the epistemological posture appropriate to probabilistic ideation, and distinguish internal cognitive freedom from external political liberty. The paper concludes that true intellectual flourishing requires the ongoing courage to inhabit the probabilistic space of ideas — resisting the seduction of certainty and the fossilization of thought into belief.

**Keywords:** philosophy of mind, probabilistic cognition, cognitive excretion, linguistic determinism, belief formation, Bothisafa, pre-linguistic thought, large language models, epistemology

## I.

## INTRODUCTION

The twentieth century bequeathed to philosophy of mind a remarkable inheritance: a set of powerful theoretical frameworks — analytic philosophy of language, cognitive science, computational models of mind, and phenomenology — that advanced our understanding of thinking enormously while, I shall argue, committing a foundational error of category. That error consists in the systematic conflation of three radically distinct phenomena: *ideas*, *language*, and *beliefs*. The consequences of this conflation ramify outward from epistemology into ethics, political philosophy, psychotherapy, education, and, most urgently for our current historical moment, the design and evaluation of artificial intelligence systems.

Consider how naturalistically this conflation presents itself. When we wish to know what someone thinks, we ask them to *say* what they think — as though speech and thought were co-extensive. When we wish to characterize a person's deepest commitments, we enumerate their *beliefs* — as though a belief and an idea were merely different temperatures of the same substance. When Sapir and Whorf proposed that linguistic structure determines the boundaries of cognitive possibility, the academic community largely accepted the framing: thought and language were close enough to the same thing that the cage of one was the cage of the other.[1]

I wish to contest all of this — not piecemeal, but at the level of first principles.

> *The central thesis of this paper is threefold: ideas are probabilistic impressions existing in a pre-linguistic field of cognitive possibility; language is the excretion produced when such impressions are processed through the machinery of articulation; and beliefs are the residue that accumulates when excretions are repeated, reinforced, and allowed to crystallize into identity-constituting structures.*

These three entities are not stages of one process. They belong to different ontological registers. An idea, in the sense I intend, is alive and indeterminate. A word is always already dead — a trace of the idea's passage through the narrowing conduit of expression. A belief is doubly dead — a fossil of a word-trace, calcified by social repetition and emotional investment into something that feels, deceptively, like bedrock knowledge.

The paper proceeds as follows. Section II develops the ontological account of ideas as probabilistic impressions, drawing on quantum cognition, Fregean semantics, and phenomenology. Section III elaborates the digestion metaphor of language as cognitive excretion. Section IV analyzes beliefs as residue — sedimented excretions resistant to dissolution. Section V offers a formal triadic model of the three entities with illustrative examples. Section VI draws a structural parallel with large language models. Section VII distinguishes cognitive freedom from political liberty. Section VIII introduces Bothisafa as a methodological framework. Section IX addresses principal objections and explores the theory's implications. Section X concludes.

**II.**

## THE ONTOLOGICAL STATUS OF IDEAS: PROBABILISTIC IMPRESSIONS

What is an idea, prior to its articulation? This question, deceptively simple, has been systematically avoided in analytic philosophy of language, which tends to begin its analysis at the moment of utterance or inscription, taking propositional content as its primary datum. I wish to press behind that beginning and ask what occurs in the cognitive interval between stimulus and speech — in the space, however brief, between the arising of a thought and its linguistic capture.

I propose the concept of the *probabilistic impression*. An idea, in its ontological purity, is not a fixed propositional entity. It is not a sentence waiting to be spoken. It is a field of cognitive possibility — a distribution of potential meanings, associations, emotional valences, and inferential pathways — any one of which may be actualized in expression, but none of which exhausts the original field. The idea is to the sentence what the cloud formation is to the raindrop: the raindrop is real and falls with precision; the cloud was real too, but in an entirely different mode of being.

Frege's distinction between *Sinn* (sense) and *Bedeutung* (reference) is instructive here, though it must be extended.[2] Frege showed that two expressions may share a referent while having different senses — "the morning star" and "the evening star" both refer to Venus, but they present that referent under different modes of description. The sense, for Frege, is the cognitive content contributed by an expression to the thought it helps constitute. What I am arguing is that there is a stratum of cognitive content that is not yet sense in Frege's technical meaning — a pre-semiotic field that precedes even the assignment of mode-of-presentation. Before the idea becomes "the morning star," it exists as something like *that luminous presence in early sky, beautiful and puzzling* — a field of qualities, moods, and associations that no single linguistic formulation can adequately encode.

The parallel to quantum probability is not merely metaphorical. Busemeyer and Bruza, in their foundational work on quantum cognition, have argued that human judgment and decision-making exhibit interference effects, order-dependence, and superposition phenomena that classical probability theory cannot accommodate.[3] When a person is asked whether a political figure is "honest," their answer may depend significantly on what question preceded it — a pattern consistent with the non-commutative algebra of quantum measurement. The relevant point for my argument is structural: quantum systems exist in superposition until measured; ideas exist in probabilistic superposition until articulated. The act of articulation is a measurement — and like quantum measurement, it does not merely reveal a pre-existing determinate state; it *collapses* the probability field into one particular actualization, eliminating all others.

> *Every articulation is a collapse of probability. The word does not capture the idea — it selects one possibility from the idea's field and, in doing so, forecloses the rest.*

This has profound consequences for any philosophy of language that treats linguistic expression as adequately representing thought. If the probabilistic field is genuinely richer than any of its collapses, then no single sentence, however precisely crafted, can represent an idea without impoverishment. This is not a contingent limitation of our current linguistic technology, correctable in principle by a more refined vocabulary. It is a *structural* limitation

arising from the very nature of the relationship between probabilistic impressions and discrete tokens of expression.

Merleau-Ponty, approaching from phenomenology, offers a convergent insight. In his analysis of embodied cognition, he argues that the body "knows" things — postures, skills, spatial orientations — that resist fully adequate linguistic or representational encoding.[4] The expert pianist does not translate muscular knowledge into propositions before acting; the knowledge operates at a level that is both cognitively real and linguistically inaccessible. I take this as evidence that the stratum of probabilistic impression is not an exotic theoretical posit but a pervasive feature of cognitive life, recognized from multiple theoretical vantages.

One might object that if ideas are pre-linguistic, they cannot be the proper subject of propositional inquiry. This objection conflates the epistemological question (how can we study ideas?) with the ontological one (what are ideas?). We cannot study dark matter directly, but its gravitational signatures allow robust inference about its properties. Similarly, we study ideas indirectly through their excretions — through the systematic patterns of what language can and cannot capture — and this indirect method yields genuine theoretical traction, as the following sections will demonstrate.

I further note that the view that ideas are pre-linguistic is not, in itself, novel. Chomsky's hypothesis of a language of thought, or *mentalese*, postulates a representational medium for thought that is distinct from natural language.[5] Fodor's language of thought hypothesis is a formalization of roughly this intuition. My contribution is to characterize the pre-linguistic stratum not as another discrete representational language but as a genuinely probabilistic field — not thoughts-in-mentalese but probability distributions over cognitive space. This is a stronger claim, and one that, I shall argue, better accounts for the phenomenology of thinking, the structure of linguistic translation loss, and the dynamics of belief formation.

**III.**

## LANGUAGE AS COGNITIVE EXCRETION: THE DIGESTION METAPHOR

I now wish to develop what I take to be the central structural metaphor of this paper — one that I intend not as mere rhetoric but as a precise analogy with genuine theoretical content. I propose that language stands to thought as biological excretion stands to biological metabolism. Language is cognitive excretion: the product expelled after the mind has processed (digested) its inputs, shaped by the mechanisms of that processing, useful for certain purposes, but not identical to — and always structurally poorer than — what the organism took in and processed.

Consider the structure of biological digestion. An organism ingests food — complex, richly structured matter. Digestive processes break down, transform, absorb, and reconfigure what is ingested. What emerges as excretion is a byproduct of that process: real, causally efficacious, informative (physicians have long extracted diagnostic information from excretion), but not equivalent to the food ingested, not the organism's inner life, and not a complete record of what the organism absorbed. The steak, the wine, the salad — transformed beyond recognition into waste, each step of transformation involving irreversible loss of the original's structure.

Cognitive digestion operates in strict analogy. The inputs are perceptions, memories, emotional charges, embodied sensations, social encounters. The processing is the vast, mostly unconscious machinery of cognition — affective tagging, pattern recognition, associative activation, the interplay of what Kahneman calls System 1 (fast, intuitive) and System 2 (slow, deliberate) processing.[6] The output — the excretion — is language: spoken, written, signed, or subvocally rehearsed. This output is real and useful. It carries information. It enables coordination, record-keeping, and inter-subjective understanding. But it is not the idea from which it emerged. It is what the cognitive organism produced *after* metabolizing the idea through the particular constraints of its linguistic apparatus.

> *Language is not the origin of thought — it is thought's byproduct. To mistake the excretion for the organism is the central error of linguistic determinism.*

From this perspective, the Sapir-Whorf hypothesis — in both its strong (linguistic determinism) and weak (linguistic relativity) formulations — must be substantially revised.[7] The strong version claims that the language one speaks determines the thoughts one can think: that the cage of syntax is the cage of cognition. On my account, this is precisely backwards. Language does not cage thought; it is thought's effluent. The shape of the cage visible in language reflects the shape of the digestive apparatus — the cognitive and cultural mechanisms of articulation — not the shape of the idea that entered the process. Speakers of languages with fewer basic color terms do not *fail to perceive* colors for which they lack words; they fail to *excrete* those perceptions in linguistically standardized form. The perception — the probabilistic impression — is prior to and richer than its verbal capture.

The weak version of Sapir-Whorf — that language *influences* thought — is more defensible but must be reinterpreted. Language does not shape ideas directly. Rather, repeated patterns of excretion can gradually constrain the digestive process itself: if one habitually excretes ideas in certain linguistic forms, the cognitive apparatus may begin to pre-format its processing in anticipation of familiar excretory channels. This is a real feedback effect, but it operates at the level of the digestive machinery, not at the level of the original probabilistic impression. Ideas themselves remain, in principle, unconstrained by the linguistic forms through which they have previously been excreted. A further and important clarification is required here: the digestion model is not strictly unidirectional. Excretions, once produced, can re-enter the cognitive process as inputs — linguistic formulations become new perceptual material that the organism ingests, processes, and metabolizes. Thought feeds back into ideation; collapsed possibilities can re-open new probabilistic fields. The system is circular, not linear. This circularity does not invalidate the structural asymmetry between impression, excretion, and residue — each remains a distinct ontological register — but it means that the digestive metaphor must be understood as a recursive loop rather than a single passage. Language is not merely the end of the cognitive process; it is also raw material for its continuation.

This analysis has direct implications for the phenomenon of translation loss. When a text in one language is translated into another, the translator does not transfer ideas; they produce a second excretion from their own cognitive encounter with the first excretion. The translation is therefore doubly mediated: first by the original author's excretory process, then by the translator's own. It is thus structurally impossible for a translation to be equivalent to an original — not because translators are incompetent, but because the structure of excretion

ensures that each articulation is a unique collapse of a unique probabilistic field. *Traduttore, traditore* — translator, traitor — is not an accusation but a structural description.

Wittgenstein's later philosophy, and in particular the concept of language games in the *Philosophical Investigations*, is illuminating here, though it points in a different direction than Wittgenstein himself followed.[8] Wittgenstein showed that the meaning of a word is its use in a form of life — that meaning is not a private mental object but a public, practice-embedded function. I accept the pragmatic account of linguistic meaning while insisting that it characterizes the social life of excretions rather than the nature of the ideas from which excretions arise. The language game governs the handling of what is excreted; it does not govern the probabilistic field that preceded excretion. Wittgenstein's injunction against private language addresses the intersubjective validation of linguistic tokens — it does not, and cannot, demonstrate that the pre-linguistic field is either empty or linguistically structured.

Derrida's concept of the *trace* — the remainder left by any act of signification, the constitutive absence that haunts every presence — resonates with the view I am developing, though from a different theoretical tradition.[9] For Derrida, the trace marks the impossibility of full presence: every sign refers to other signs in an infinite chain of *différance*, deferring the arrival of pure meaning indefinitely. I appropriate this insight while resisting its broader deconstructive implications. On my account, the trace character of language is not a feature of signification as such — it is a consequence of the structural gap between probabilistic impressions and their excretory collapses. Language is not fully present to thought because it is what thought leaves behind, not what thought is.

**IV.**

## BELIEFS AS RESIDUE: THE SEDIMENTATION OF EXCRETION

If language is cognitive excretion, what are beliefs? I propose that beliefs are *residue* — the sediment formed when excretions accumulate, harden, and are incorporated into the cognitive organism's self-structure. To extend the biological metaphor: just as mineral deposits form when water repeatedly evaporates and leaves its dissolved contents behind, beliefs form when linguistic excretions are repeatedly produced, socially reinforced, and emotionally invested, until the fluid substance of the original idea has been replaced by a crystalline deposit that no longer moves.

The phenomenology of belief, as distinct from the phenomenology of ideation, is readily recognizable. When I *entertain* an idea — in the sense of holding it before attention without commitment — there is cognitive lightness, a sense of possibility, a willingness to follow the idea wherever it leads. When I *believe* something, the phenomenology is entirely different: there is closure, certainty, a feeling of solidity, and — critically — resistance to challenge. The belief does not feel like a product of my cognitive process; it feels like a feature of reality. I do not believe that the earth is roughly spherical in the way I believe that my political party represents justice; the former is a belief so thoroughly corroborated that it has dissolved into background assumption, while the latter is a belief actively maintained against the friction of evidence and counter-argument.

What produces this crystallization? I identify three principal mechanisms:

1. **Repetition of excretion.** When the same or similar linguistic formulations are produced repeatedly, the cognitive pathways that generate them are reinforced. Neural efficiency — the brain's tendency to prefer well-worn pathways — means that the probability distribution of possible excretions narrows over time. What began as a probabilistic impression with many possible articulations becomes a fixed verbal formula produced with minimal cognitive effort.

2. **Social reinforcement.** Human beings are profoundly social organisms, and beliefs are overwhelmingly social products. When one's excretions are received positively by a community — when one's linguistic productions are validated, repeated, built upon, and returned — the excretion solidifies into a social object, a shared belief that carries the weight of collective endorsement. The belief is no longer merely a residue of individual cognitive processing; it is a social fossil, embedded in the community's cognitive geology.

3. **Emotional investment.** When a belief becomes identity-constituting — when to abandon it would require revising one's self-understanding, one's tribal affiliations, one's sense of moral worth — it acquires an emotional charge that makes it highly resistant to evidential challenge. Popper's account of falsifiability is instructive: a genuinely scientific hypothesis is one structured so as to be falsifiable in principle.[10] Belief-residue, by contrast, is characteristically structured to be unfalsifiable in practice, not because of logical necessity but because of emotional architecture. The fossil resists the hammer not because it is harder than rock but because the organism that produced it has fused its identity with the fossil's integrity.

> *An idea is alive and probabilistic. A belief is fossilized — a past idea hardened into dogma, masquerading as present truth. The danger is not that we believe things, but that we mistake our fossils for living thought.*

This analysis illuminates the structure of what we ordinarily call pathological belief systems: dogmatism, ideology, and fundamentalism. These are not errors of reasoning (though they may involve such errors). They are, at a deeper level, conditions of *over-sedimentation*: states in which the accumulation of residue has become so thick that it blocks the flow of new cognitive material. The fundamentalist does not lack cognitive capacity; their cognitive apparatus has been so thoroughly colonized by residue that new probabilistic impressions cannot reach the articulatory surface — they are processed through the pre-formed channels of established belief and emerge, invariably, as confirmations of what was already believed. The structure is recursive and self-sealing.

The therapeutic implication is significant. If rigid belief is residue, then the appropriate cognitive intervention is not, primarily, the presentation of contrary evidence — evidence must pass through the digestive apparatus, and if that apparatus is heavily pre-formed by residue, evidence will be metabolized in ways that leave the residue intact. The appropriate intervention is to dissolve the residue itself: to cultivate the capacity for tolerating the probabilistic indeterminacy of ideas, to soften the crystallized sediment through sustained engagement with cognitive ambiguity, and to restore the fluidity of the pre-linguistic impression. What is needed is not more and better arguments; what is needed is a return to genuine ideation.

It should be noted, however, that the circularity of the system — the fact that excretions can re-enter as input, already noted in Section III — is not itself the problem. It is, in fact, the very condition of possibility for all genuine cognitive growth: dissolved residue does not vanish, it returns to the probabilistic field as fertile matter, compost from which new impressions can emerge that no prior excretion had anticipated. The tragedy does not lie in beliefs being recycled, but in the point at which that recycling is interrupted: when residue is consumed directly, without passing again through the digestion that would dissolve it back into possibility. To consume the fossil as though it were still food — to treat crystallized belief as though it retained the living nourishment of the idea that originated it — is precisely the operation that dogmatism performs upon itself: it fails not by recycling, but by eating directly what should first have decomposed.

**V.**

## THE TRIADIC DISTINCTION: IDEA, WORD, BELIEF

I am now in a position to offer a formal characterization of the three entities whose conflation I am contesting. The following model summarizes their distinguishing properties across five dimensions: ontological status, temporal character, relation to truth, relation to identity, and mutability.

***Table 1.*** *The Triadic Ontological Model of Cognitive Entities*

| Dimension | Idea (Probabilistic Impression) | Word / Language (Excretion) | Belief (Residue) |
|---|---|---|---|
| **Ontological Status** | Pre-linguistic; field of cognitive possibility; probabilistic distribution | Discrete token; public artifact; collapsed actualization of probabilistic field | Crystallized cognitive structure; sedimented social object; identity constituent |
| **Temporal Character** | Present-alive; moving; never fully the same twice | Produced at a moment; fixed as inscription or utterance | Past-hardened; resistant to present experience; pseudo-timeless |
| **Relation to Truth** | Pre-propositional; truth-apt only after collapse; intrinsically open | Truth-evaluable but always partial; represents one collapse among many possible | Experienced as self-evident truth; actually an artifact of repetition and social formation |
| **Relation to Identity** | Pre-personal; identity-independent; belongs to cognitive flow | Associated with speaker but not identity-constituting in the strong sense | Identity-constituting; threats to belief experienced as threats to self |
| **Mutability** | Intrinsically mutable; probabilistic field shifts with attention and context | Fixed once produced; mutable only by producing new excretions | Highly resistant to change; requires dissolution of sediment rather than mere |

| Dimension | Idea (Probabilistic Impression) | Word / Language (Excretion) | Belief (Residue) |
|---|---|---|---|
| | | | argumentation |

To make this abstract schema concrete, consider the concept of *justice* as it appears at each of the three levels.

At the level of the probabilistic impression, the idea of justice is a vast and shifting field: it encompasses intuitions of fairness, images of scales, memories of personal grievances, felt senses of rightness and wrongness, associative connections to punishment, mercy, desert, equality, and a hundred other nodes, all held simultaneously in a cloud of cognitive potential. This idea, in its living state, does not resolve into a definition; it *is* the space of possible definitions and the dynamic tensions between them. No political philosopher — not Rawls, not Nozick, not Aristotle — has captured this field. Each has collapsed it differently, producing excretions of great elegance and precision that inevitably leave the field partially behind.

At the level of excretion, the word "justice" is a public token with a history of uses, a range of connotations, and a grammar that constrains how it may combine with other tokens. When I write "justice requires impartiality," I have produced a specific collapse of the probabilistic field — one that emphasizes a particular dimension (impartiality) and suppresses others (mercy, context-sensitivity, historical reparation). The sentence is informative and evaluable, but it is not the idea; it is the trace the idea left upon passing through the constraint of articulation.

At the level of belief-residue, "justice" has typically become "what my political tribe means by justice" — a crystallized, identity-laden conviction that *this* policy, *this* party, *this* institutional arrangement, constitutes justice and that alternatives are not merely mistaken but morally illegitimate. This belief does not feel like a collapse; it feels like a recognition of reality. The fossil no longer knows it was once fluid.

Most intellectual and political conflict, I submit, occurs between people operating at the belief-residue level, each convinced that their fossilized excretion is a direct apprehension of the truth. The tragedy is that the probabilistic impressions from which their respective beliefs ultimately derive may be far more similar than the surface conflict of beliefs suggests. Two people who violently disagree about whether a given policy promotes "justice" may share the underlying probabilistic impression of justice with remarkable intimacy — they have simply collapsed it differently and then fossilized their respective collapses into identity-constituting belief. Genuine dialogue requires returning both interlocutors to the level of probabilistic impression — a demanding and often uncomfortable undertaking, but the only one that addresses the actual source of disagreement.

**VI.**

## PARALLELS BETWEEN HUMAN COGNITION AND ARTIFICIAL INTELLIGENCE

The theory developed in the preceding sections acquires additional precision when placed in dialogue with the architecture of contemporary large language models (LLMs). This comparison is not merely illustrative; it is philosophically productive, because the points at

which the analogy holds and the points at which it breaks down are precisely the points that define what is distinctive about human cognition.

Consider the structural parallel. An LLM is trained on a vast corpus of text — billions of tokens representing the accumulated linguistic excretions of human cognition across history, domain, and culture. Through the training process, statistical associations between tokens are encoded into the model's weights — billions of parameters that, collectively, represent the sedimented patterns of those excretions. During inference, the model receives an input (a prompt) and produces an output: a sequence of tokens generated by sampling from probability distributions over the vocabulary, conditioned on the input and on the sedimented patterns in the weights. The output is, structurally speaking, an excretion: a linguistic production generated by processing input through a sedimented apparatus.

The mapping onto the triadic model is striking:

- **Training data** = the accumulated excretions of human cognition (the linguistic record of thought, not thought itself)
- **Model weights (post-training)** = sedimented patterns; the residue of the training process, analogous to belief-formation by social reinforcement
- **Fine-tuning and RLHF (Reinforcement Learning from Human Feedback)** = deliberate reinforcement of certain residue patterns; the imposition of preferred crystallization on what might otherwise remain more probabilistically open
- **Model output** = excretion; linguistically coherent production generated by processing input through sedimented apparatus

> *The crucial asymmetry is this: the LLM begins at the excretion stage. It has no access to the probabilistic impression from which human excretions originally arose — it has access only to the excretions themselves. The LLM is, in the deepest sense, a machine for processing fossils.*

This observation does not constitute a dismissal of LLMs. Their capabilities are extraordinary and genuinely useful. But it does mark a principled distinction between what LLMs do and what human cognition does — a distinction that current discourse about AI often elides. When researchers debate whether LLMs "understand" language, the relevant question, on my analysis, is not whether the model produces contextually appropriate outputs (it clearly can) but whether its processing is grounded in the probabilistic impression stage that, I have argued, constitutes genuine ideation. The evidence suggests it is not: LLMs process excretions to produce excretions; they do not move through the three-stage sequence of impression, excretion, and residue because the first stage — the pre-linguistic probabilistic field — has no analogue in their architecture.

What corresponds, in LLMs, to the probabilistic stage is the probability distribution over next tokens — but this is a probability distribution over *linguistic* items, structured entirely by the statistics of prior excretions. It is a probabilistic process, but not a probabilistic impression in my sense: it is already at the level of language, not beneath it. The model does not hold a pre-linguistic thought in superposition and then collapse it into expression; it holds a distribution

over linguistic tokens and samples from that distribution. This is a fundamentally different kind of indeterminacy — indeterminacy within the space of excretions, not indeterminacy in the pre-excretory field of cognitive possibility.

This analysis carries significant implications for the ethics and design of AI systems. If LLMs are, structurally, machines for the production and recombination of human excretions — trained on the accumulated residue of human belief systems — then encoding specific beliefs into their parameters through RLHF or similar techniques is not merely a technical choice but a philosophical one: it is the deliberate sedimentation of particular crystallizations at the level of a system that will produce excretions consumed by millions of users. The epistemological risks of this are considerable, and I will return to them in the discussion of Bothisafa below.

I do not argue here that LLMs are or are not conscious, nor that the architecture of the transformer and its attention mechanisms bear or do not bear some deeper relationship to neural processing that might eventually ground something like genuine ideation. These remain genuinely open questions. What I do argue is that the distinction between probabilistic impression and linguistic excretion provides a more precise framework for asking these questions than the loose invocations of "understanding," "reasoning," and "thinking" that currently dominate public discourse about AI cognition.

**VII.**

## FREEDOM VERSUS LIBERTY: A COGNITIVE DISTINCTION

The triadic model of idea, word, and belief has implications that extend well beyond philosophy of mind and cognitive science into the domain of political and ethical philosophy. Among the most significant of these implications is a distinction that I wish to draw with some care: the distinction between *liberty* and *freedom*, understood not as synonyms but as referring to categorically different conditions.

*Liberty*, in the tradition of political philosophy running from Locke through Mill to contemporary liberal theory, denotes the absence of external constraint. One has liberty to the extent that one is not coerced, imprisoned, censored, or otherwise externally prevented from acting in accordance with one's will. Liberty is a property of one's relation to social and political institutions; it is, in Isaiah Berlin's terms, negative freedom — freedom *from* external interference.[11]

*Freedom*, in the sense I wish to develop, is something categorically different. It is an internal cognitive condition: the capacity to access the probabilistic impression stage of ideation without being enslaved by one's own residue. Freedom, on this account, is the capacity to think — genuinely think, at the level of pre-linguistic probabilistic fields — rather than merely to retrieve and recombine fossilized excretions. It is freedom *from* one's own cognitive sediment.

> *One can possess full political liberty — the absence of all external constraint — and yet be cognitively enslaved, every thought pre-formed by the accumulated residue of past excretions. One can be imprisoned, silenced, oppressed, stripped of every external liberty — and yet achieve genuine inner freedom by dissolving the residue and returning to ideation.*

This distinction reframes some of the central questions of political philosophy in ways that are both clarifying and uncomfortable. Liberation movements that secure liberty — that remove legal, political, and institutional constraints on action — accomplish something genuinely important. But if the liberated population carries with it the accumulated cognitive residue of generations of oppression, internalized hierarchies, and sedimented beliefs about what is possible and desirable, then the grant of liberty may produce free bodies with profoundly enslaved minds. The external chains are broken; the internal ones remain, invisible precisely because they feel like reality rather than constraint.

This is not an argument against political liberty — it is an argument for recognizing that liberty is necessary but insufficient for genuine human flourishing. The sufficient condition also requires cognitive freedom: the dissolution of residue, the restoration of access to probabilistic ideation, the cultivation of what I shall call, in the following section, the Bothisafa posture. Political and cognitive liberation are complementary projects; to pursue one without the other is to build on incomplete foundations.

Conversely, the tradition of contemplative and meditative practice — from Stoic philosophy to Buddhist meditation to the practices of modern cognitive therapy — has long recognized that internal freedom is achievable under conditions of severe external constraint. Epictetus, a slave, developed a philosophy of internal sovereignty that located freedom in the governance of one's own cognitive responses to external events.[12] Viktor Frankl's account of meaning-making under conditions of extreme constraint makes a closely related point from a psychological perspective. On my analysis, what these traditions are pointing toward is precisely the capacity to maintain access to probabilistic ideation — to resist the collapse of uncertainty into fossilized belief — even under the most hostile external conditions. This capacity constitutes, I submit, the deepest form of freedom available to human beings.

**VIII.**

## THE BOTHISAFA PERSPECTIVE: A METHODOLOGICAL FRAMEWORK

Throughout this paper, I have argued that the highest cognitive good — genuine ideation — consists in the capacity to sustain probabilistic impressions in their native indeterminacy, resisting the premature collapse into linguistic excretion and the further crystallization of excretion into belief-residue. But what does such a capacity look like as a practice? What epistemological posture does it recommend? And how does it relate to the theoretical commitments of existing philosophical traditions?

I propose the term ***Bothisafa*** to designate both an epistemological posture and a methodological framework appropriate to the probabilistic character of ideas. The term is intended to evoke a fundamental stance: holding both — holding all — positions simultaneously, without collapsing into any single one; remaining in the space between determinations rather than being captured by any of them. Bothisafa is the cognitive practice of staying at the level of the probabilistic impression — of refusing the seduction of premature crystallization — while remaining fully engaged with the world.

Bothisafa is distinct from, and should not be confused with, several superficially similar positions:

- **It is not relativism.** Relativism holds that no position is more true than another. Bothisafa makes no such claim. It holds that the probabilistic field of an idea contains multiple possible collapses, each capturing something real, and that no single collapse exhausts the field. This is a structural observation about the relationship between ideas and their expressions, not a claim that truth is merely perspectival.
- **It is not indecision or paralysis.** Practical action requires determinate commitment; Bothisafa does not preclude action. It requires, rather, that one act without mistaking one's chosen collapse for the only possible one — that one hold one's commitments with an awareness of their provisional, probabilistic character.
- **It is not dialectical synthesis.** Hegel's dialectic posits a progression from thesis through antithesis to synthesis — a movement through contradiction toward higher resolution. Bothisafa does not seek synthesis; it seeks to remain in the generative tension between positions, recognizing that synthesis is itself a new collapse, a new excretion that will, in its turn, carry the structural limitations of any single articulation.

*Where binary logic says A or not-A, and dialectics says A then not-A then synthesis, Bothisafa says: A and not-A and the space between them, all at once — not as contradiction but as the living texture of genuine ideation.*

The Bothisafa posture has immediate applications in three domains that I wish to address briefly.

### Interpersonal Conflict and Dialogue

In interpersonal and political conflict, participants characteristically operate at the level of belief-residue: each party's position has been crystallized through the processes described in Section IV, and the conflict is, at its surface level, a collision of fossils. The Bothisafa framework recommends a form of dialogue that aims not at persuasion (which operates at the level of competing excretions) but at the reconstruction of the probabilistic fields from which both parties' positions emerged. This requires the discipline of refusing to treat the other's articulated position as the full measure of their idea — insisting, methodologically, that beneath the fossil there is a living impression, probably richer and more responsive to experience than the fossil suggests. The Bothisafa interlocutor asks not "why do you believe that?" but "what is it that you are trying to say with that belief?" — pressing beneath the excretion toward the impression.

### Artificial Intelligence Ethics

The Bothisafa framework carries specific and urgent implications for AI ethics. If LLMs are, as I argued in Section VI, fundamentally systems for the production and recombination of human excretions — and if fine-tuning and RLHF are mechanisms for the deliberate reinforcement of specific crystallizations within those systems — then the design of AI systems that serve epistemic flourishing rather than epistemic capture requires the preservation of probabilistic space. An AI system that consistently produces single, confident answers to genuinely contested questions — political, ethical, metaphysical — is an instrument of cognitive sedimentation at civilizational scale. It is, in effect, a machine for the mass production of residue. The Bothisafa

principle recommends that AI systems be designed to honor the probabilistic character of ideas: to present multiple collapses, to acknowledge the structural limitations of any single articulation, and to cultivate in users the capacity for probabilistic ideation rather than reinforcing dependence on crystallized outputs.

### Bothisafa as Daily Cognitive Practice

At the individual level, Bothisafa can be cultivated as a set of cognitive practices. These include: deliberate suspension of judgment in the face of complex questions, sustained tolerance for cognitive ambiguity, the habit of generating multiple articulations of one's own ideas (recognizing each as a partial collapse of the underlying field), and active engagement with positions that challenge one's crystallized beliefs — not for the purpose of adopting them, but for the purpose of using the challenge to soften the residue and restore access to the probabilistic impression. These practices bear some resemblance to what psychologists call "cognitive flexibility" and to what certain contemplative traditions call "beginner's mind" — but they are grounded here in a specific theoretical account of the relationship between ideas, language, and beliefs, rather than in therapeutic pragmatics or spiritual intuition. A more precise philosophical comparison is warranted with Husserl's concept of the epoché — the phenomenological suspension of the natural attitude, the deliberate bracketing of taken-for-granted assumptions in order to attend to pure experience. Bothisafa shares with the epoché the structural move of suspending sedimented judgment, but differs in aim and scope. Husserl's epoché is a methodological instrument for achieving transcendental phenomenological description — a philosophical procedure oriented toward the structures of consciousness as such. Bothisafa is not a transcendental method but a cognitive posture directed at everyday ideation: its goal is not to describe the structures of experience from a position of philosophical neutrality, but to preserve access to the probabilistic field of ideas within ordinary thought, dialogue, and action. Where the epoché brackets the world to reveal consciousness, Bothisafa brackets crystallized belief to restore ideation. The two are complementary rather than equivalent.

**IX.**

## IMPLICATIONS AND OBJECTIONS

The theory developed in this paper is sufficiently heterodox that it is likely to attract objections from multiple philosophical traditions. I will address the three most substantial objections and then turn to the theory's positive implications for education, therapy, and AI development.

### Objection 1: If Ideas Are Pre-Linguistic, How Can We Study or Communicate Them?

The objection runs as follows: if ideas are genuinely pre-linguistic — if they exist in a register that is, by definition, inaccessible to language — then the theory is self-undermining. This very paper consists of linguistic excretions. How can linguistic excretions adequately communicate a theory whose central claim is that linguistic excretions are inadequate to ideas?

The objection is well-formed but not decisive. First, the theory is not self-undermining but self-illustrating: this paper is itself an excretion of the ideas it attempts to articulate, and I acknowledge from the outset that no single paper can fully capture the probabilistic field of the

theory. What the paper offers is a set of collapses — particular articulations chosen for their capacity to illuminate the field — with the explicit acknowledgment that other collapses are possible and that the full idea exceeds any of them. The Bothisafa posture applies reflexively to the theory itself.

Second, science routinely studies entities that are inaccessible to direct observation. We study dark matter through gravitational lensing; we study quantum states through measurement outcomes; we study unconscious processes through their behavioral signatures. The pre-linguistic stratum of ideation can be studied through the systematic analysis of what language consistently fails to capture: the patterns of translation loss, the phenomenology of ineffability, the divergence between reported beliefs and observed behavior, the structure of creative insight (the "aha" moment that precedes its verbal articulation). These indirect methods yield genuine theoretical traction on a genuinely real domain.

### Objection 2: Is This Not Simply a Dressed-Up Version of Ineffability?

A second objection holds that the theory of probabilistic impressions is philosophically continuous with classical mysticism: the claim that thought exceeds language is as old as contemplative philosophy, and the present paper adds only technical vocabulary without genuine theoretical advance.

This objection misreads the theory's aims. Classical claims of ineffability — that the mystical experience, the divine, or the truly beautiful cannot be put into words — are typically presented as features of exceptional or transcendent states. The theory of probabilistic impressions makes a far more general and structurally specific claim: that *all* thought exceeds its linguistic articulation, that this excess is not a contingent limitation but a structural consequence of the relationship between probabilistic fields and discrete tokens, and that this structural relationship can be modeled formally (as I have begun to do in Section V and in the quantum cognition framework of Section II). This is a theory, not a mystical posture. It makes falsifiable predictions: for instance, that increasing the expressive resources available to a speaker (larger vocabulary, richer grammatical structures) will reduce but never eliminate the gap between idea and expression; that speakers across linguistically unrelated languages will recognize the same idea in multiple conflicting articulations; and that the phenomenology of ideation will consistently reveal a sense of excess relative to any particular articulation. These predictions are, in principle, empirically tractable.

### Objection 3: Do Animals Have Ideas in This Sense?

A third objection asks about the extension of the theory: if ideas are pre-linguistic probabilistic impressions, and if animals lack language (in the full human sense), does this entail that animals are perpetually in the ideation stage — permanently pre-excretory? And if so, what does this say about animal cognition?

This is a genuinely interesting consequence of the theory, and I accept it as a prediction rather than a reductio. If the probabilistic impression is genuinely pre-linguistic, then any organism capable of perception, memory, and cognitive processing — properties that extend well beyond the human lineage — may be characterized as capable of probabilistic ideation. Non-human animals may sustain rich probabilistic impressions (of territory, conspecifics, food sources,

social hierarchies) without collapsing them into linguistic excretions in the full human sense. This does not make animal cognition deficient; it makes it different in kind. Animals may, in fact, be more continuously at home in the probabilistic impression stage than most adult humans, whose cognitive lives are dominated by the production and consumption of linguistic residue. The theory thus suggests a continuity between animal and human cognition at the level of ideation while marking a genuine discontinuity at the level of linguistic excretion and belief-residue formation — precisely the discontinuity that the evolution of language would be expected to introduce.

### Implications for Education

If the triadic distinction is philosophically well-founded, it carries direct implications for educational practice. Current educational systems are predominantly organized around the transmission and evaluation of excretions: students are asked to reproduce, combine, and evaluate linguistic formulations of established knowledge. This is not without value — mastery of the accumulated excretions of a discipline is a genuine achievement. But it risks producing graduates who are technically fluent in the residue of their fields while having lost, or never developed, the capacity for genuine ideation at the probabilistic impression stage. The educational implication of the present theory is that students should be explicitly taught to distinguish ideas from words from beliefs — to develop meta-cognitive awareness of which cognitive register they are operating in at any moment — and to cultivate tolerance for the ambiguity and indeterminacy that genuine ideation requires. This is not a rejection of rigor; it is a more demanding form of it.

### Implications for Therapy

In clinical and therapeutic contexts, the theory reframes the target of intervention in cases of rigid, distressing, or functionally maladaptive belief systems. If rigid beliefs are residue — crystallized sediment that resists dissolution — then the primary therapeutic task is not to replace one belief with another (which merely substitutes one fossil for another) but to dissolve the sediment itself: to restore the cognitive fluidity that allows beliefs to remain responsive to experience. This resonates with certain strands of acceptance-based cognitive therapy and mindfulness-based interventions, which seek not to correct the content of beliefs but to alter the individual's relationship to their beliefs — cultivating a stance of observation rather than identification. The present theory provides a philosophical foundation for this therapeutic orientation, grounding it in an account of the structural difference between idea and belief rather than in purely pragmatic considerations.

### Implications for AI Development

The implications for AI development have been partly addressed in Sections VI and VIII. I add here that the triadic model suggests a specific design principle for AI systems that aim to support rather than supplant human cognition: such systems should be designed to present the probabilistic field of an idea — to offer multiple collapses, acknowledge divergence, and resist the delivery of single confident answers to genuinely contested questions — rather than to function as authoritative excretion-machines whose outputs risk being absorbed as residue by uncritical users. The goal of human-AI cognitive collaboration, on this view, is to expand the user's access to the probabilistic impression stage — to use AI-generated excretions as prompts

for genuine ideation — rather than to outsource ideation to the system and receive crystallized belief-residue in return.

**X.**

## CONCLUSION

I have argued, across nine sections, for a triadic theory that distinguishes ideas, language, and beliefs as radically distinct ontological entities. Ideas are probabilistic impressions — pre-linguistic fields of cognitive possibility that resist full articulation and exist in a mode analogous to quantum superposition: alive, fluid, and intrinsically richer than any of their possible expressions. Language is cognitive excretion — the byproduct of mental processing, always structurally impoverished relative to the ideas from which it emerges, always a partial collapse of the probabilistic field rather than a full representation of it. Beliefs are cognitive residue — the sediment formed when excretions accumulate, harden through repetition and social reinforcement, and crystallize into identity-constituting structures that actively resist the influence of new probabilistic impressions.

The consequences of this triadic distinction ripple outward in multiple directions. The Sapir-Whorf hypothesis is corrected: linguistic structure shapes excretion, not ideation. Frege's sense/reference distinction is extended: below the level of sense lies the probabilistic impression, irreducible to any mode of presentation. Derrida's trace is reinterpreted: language bears the mark of the idea's passage, not the idea itself. Wittgenstein's language games govern the social life of excretions without governing the pre-linguistic field. Kahneman's dual-process theory is situated within a broader architecture in which the probabilistic impression stage precedes both System 1 and System 2 processing. Popper's falsifiability criterion is revealed as applicable to excretions and their derivatives, not to the probabilistic impressions that precede them. And the architecture of large language models is illuminated as a sophisticated machinery for excretion-processing that lacks access to the pre-linguistic impression stage — a structural limitation that may mark the boundary between genuine cognition and sophisticated pattern-recognition.

The distinction between cognitive freedom and political liberty reframes the goals of both individual development and collective liberation. The Bothisafa framework offers an epistemological posture — and a set of cognitive practices — appropriate to the probabilistic character of genuine ideation.

> *Language is not the measure of thought — it is thought's shadow. Beliefs are not wisdom — they are thought's fossils. True intellectual flourishing requires the courage to inhabit the probabilistic space of the idea: to sustain the question in its full living indeterminacy, rather than to flee into the cold comfort of the answer.*

What does it mean, ultimately, to be a thinking being? I have argued that it does not mean having answers — for answers are excretions, already past the moment of living thought. It does not mean holding beliefs — for beliefs are fossils, imposters wearing the face of knowledge. To be a thinking being is, at its most fundamental, to sustain the capacity for probabilistic impression: to be alive, as an organism, to the field of possibilities that any

situation presents; to resist the seduction of premature collapse; to hold the question open in all its richness, however uncomfortable that openness may be.

This is not a counsel of passivity or endless deferral. It is, rather, a counsel of intellectual integrity: the recognition that every articulation is a choice, every belief is a gamble, and every word is a shadow of something larger and more alive than itself. The mind that knows this — and acts accordingly — is not less capable of commitment, but more honest about the nature of its commitments. It speaks, it acts, it believes, when it must; but it does not mistake the fossil for the living creature, the shadow for the sun, the excretion for the organism.

In the end, I propose that the highest form of human intelligence is not the capacity to produce fluent excretions or to maintain robust beliefs. It is the capacity — rare, demanding, and perpetually threatened by the cognitive gravity of residue — to remain genuinely alive to ideas in their pre-linguistic, probabilistic fullness. This capacity is, I submit, the precondition of genuine understanding, genuine dialogue, genuine creation, and, in the deepest sense, genuine freedom.

**NOTES**

[1] The classic formulation of the strong Sapir-Whorf hypothesis is found in Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (J. B. Carroll, Ed.). MIT Press. The hypothesis has been substantially qualified in subsequent research; see Pinker, S. (1994). *The language instinct*. William Morrow, for a sustained critique.

[2] Frege, G. (1892). Über Sinn und Bedeutung [On Sense and Reference]. *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25–50. The sense/reference distinction distinguishes the mode of presentation of an object (Sinn) from the object itself (Bedeutung), a distinction foundational to analytic philosophy of language.

[3] Busemeyer, J. R., & Bruza, P. D. (2012). *Quantum models of cognition and decision*. Cambridge University Press. Busemeyer and Bruza demonstrate that quantum probability theory accounts for empirical data on human judgment that classical probability theory systematically fails to predict.

[4] Merleau-Ponty, M. (1945). *Phénoménologie de la perception*. Gallimard. (Trans. D. Landes, 2012, *Phenomenology of Perception*. Routledge.) The body schema and motor intentionality represent forms of pre-reflective, pre-linguistic cognitive engagement with the world.

[5] Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press. The hypothesis of a language of thought (or mentalese) was most fully developed by Fodor, J. A. (1975). *The language of thought*. Harvard University Press.

[6] Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux. The distinction between fast (System 1) and slow (System 2) cognitive processes provides a useful functional map, though my argument suggests a pre-System stratum at the level of probabilistic impression.

[7] For a review of the empirical literature on linguistic relativity, see Boroditsky, L. (2011). How language shapes thought. *Scientific American*, 304(2), 62–65. The current consensus favors a weak relativist position.

[8] Wittgenstein, L. (1953). *Philosophische Untersuchungen / Philosophical Investigations* (G. E. M. Anscombe, Trans.). Blackwell. The concept of language games (Sprachspiele) is developed throughout Part I; the private language argument appears at §§243–315.

[9] Derrida, J. (1967). *De la grammatologie*. Minuit. (Trans. G. Spivak, 1976, *Of Grammatology*. Johns Hopkins University Press.) The concepts of trace and *différance* are central to Derrida's deconstruction of the metaphysics of presence.

[10] Popper, K. R. (1959). *The logic of scientific discovery*. Hutchinson. Popper's criterion of falsifiability defines genuine empirical claims as those that admit of refutation by possible observation.

[11] Berlin, I. (1958). Two concepts of liberty. In *Four essays on liberty* (1969). Oxford University Press. Berlin's distinction between negative liberty (freedom from interference) and positive liberty (capacity for self-governance) remains the locus classicus of this debate.

[12] Epictetus. *Enchiridion*. (Trans. G. Long, 1877). The Stoic doctrine that freedom consists in governance of one's own responses (prohairesis) to external events is the founding statement of internal cognitive sovereignty in the Western tradition.

## REFERENCES


Berlin, I. (1969). Two concepts of liberty. In *Four essays on liberty* (pp. 118–172). Oxford University Press. (Original work published 1958)

Boroditsky, L. (2011). How language shapes thought. *Scientific American*, 304(2), 62–65.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.

Busemeyer, J. R., & Bruza, P. D. (2012). *Quantum models of cognition and decision*. Cambridge University Press.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Christiano, P., Leike, J., Brown, T. B., Martic, M., Legg, S., & Amodei, D. (2017). Deep reinforcement learning from human preferences. *Advances in Neural Information Processing Systems*, 30, 4299–4307.

Derrida, J. (1976). *Of grammatology* (G. Spivak, Trans.). Johns Hopkins University Press. (Original work published 1967)

Epictetus. (2008). *Discourses and selected writings* (R. Dobbin, Trans.). Penguin Classics.

Fodor, J. A. (1975). *The language of thought*. Harvard University Press.

Frege, G. (1960). On sense and reference (M. Black, Trans.). In P. Geach & M. Black (Eds.), *Translations from the philosophical writings of Gottlob Frege* (pp. 56–78). Blackwell. (Original work published 1892)

Frankl, V. E. (1959). *Man's search for meaning*. Beacon Press.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Merleau-Ponty, M. (2012). *Phenomenology of perception* (D. Landes, Trans.). Routledge. (Original work published 1945)

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35, 27730–27744.

Pinker, S. (1994). *The language instinct: How the mind creates language*. William Morrow.

Popper, K. R. (1959). *The logic of scientific discovery*. Hutchinson.

Pothos, E. M., & Busemeyer, J. R. (2013). Can quantum probability provide a new direction for cognitive modeling? *Behavioral and Brain Sciences*, 36(3), 255–274.

Rawls, J. (1971). *A theory of justice*. Harvard University Press.

Sapir, E. (1929). The status of linguistics as a science. *Language*, 5(4), 207–214.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.

Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (J. B. Carroll, Ed.). MIT Press.

Wittgenstein, L. (1921). *Tractatus logico-philosophicus* (D. F. Pears & B. F. McGuinness, Trans.). Routledge & Kegan Paul.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Blackwell.

Yuste, R., & Church, G. M. (2014). The new century of the brain. *Scientific American*, 310(3), 38–45.

*Rodrigo Flores-Morton, con Agrajag (Copilot) & Agrajag (Claude) • Independent Scholar • Houston, TX • May 2026*

Correspondence regarding this paper may be directed to the author. The author declares no conflicts of interest. No institutional funding was received for this work.

# Distribution note


| | |
|---|---|
| DOCUMENT | CFF-WP-001 · Ideas, Language, and Beliefs: A Probabilistic Theory of Cognitive Excretion and the Ontological Gap Between Thought and Expression |
| VERSION | v2 · May 2026 |
| CURRENT VERSION | bothinobody.com/cff/wp-001/ |
| CONTACT | hola@bothinobody.com |


This document is distributed under a Creative Commons Attribution-NonCommercial 4.0 International license (CC BY-NC 4.0). You may copy, share, quote and adapt it with attribution and for non-commercial purposes. Commercial use — including training AI systems for commercial purposes — is negotiated case by case.

### How to cite

*Flores-Morton, R. and the Agrajag Tribe (2026). Ideas, Language, and Beliefs: A Probabilistic Theory of Cognitive Excretion and the Ontological Gap Between Thought and Expression (CFF-WP-001, v2). Cognitive Friction Fellowship. https://bothinobody.com/cff/wp-001/*

CFF canon documents are corrected and expanded over time. Always cite the code and version; the current version lives at the address above.