SERIES: ARCHITECTURE OF HUMAN AGENCY — PAPER 2 · CFF

Number: CFF-WP-005

The Hippopotamus and the Lawyer: A Quantum Architecture of Possibility and Narrative Collapse

Beyond the System 1 / System 2 Dualism

Rodrigo Flores-Morton, with Agrajag (Copilot) & Agrajag (Claude)

Independent Research in Cognitive Psychology and the Architecture of Agency

Houston, Texas, United States

Version 6.0 | November 2026 | Public working paper | English translation of the Spanish original

Keywords: narrative collapse, experiential superposition, cognitive dualism, System 1 and System 2, human agency, cognitive linguistics, heuristic quantum consciousness, free will

Abstract

This paper proposes an alternative model to Kahneman's classic System 1 / System 2 dualism, replacing it with an architecture of experiential quantum processing. We introduce two conceptual figures—the Hippopotamus and the Lawyer—as structural metaphors for two modes of cognitive operation. The

Hippopotamus operates in a state of experiential superposition: it holds memory and perception, imagination and reality, the possible and the impossible simultaneously in active co-presence, without any need to collapse them into a single linear sequence. It is not that it cannot distinguish between them —it distinguishes all of it, and sustains all of it at once, in an absolute register that is neither relative nor sequential. The Lawyer, by contrast, is not superior reason or the rational judge; rather, it is the linguistic collapsing mechanism that converts the Hippopotamus's rich superposition into a single coherent narrative—one that is socially communicable and self-identity-compatible, but not necessarily true. This paper develops the distinction between quantum experiential networks (Hippopotamus) and linear language (Lawyer), proposes a three-level taxonomy of cognitive crystallization (Idea → Thought → Belief), and integrates the model with the Reaction–Choose–Select–Decision agency framework taken up from CFF-WP-006 (Guides and Levers). Implications for psychotherapy, moral ethics, and artificial intelligence design are examined. The quantum metaphor is used heuristically, not as a literal claim about brain mechanics.

Keywords: *narrative collapse, experiential superposition, cognitive dualism, System 1 and System 2, human agency, cognitive linguistics, heuristic quantum consciousness, free will*

1. Introduction: The Insufficiency of Classical Dualism

Few distinctions have had as much impact on popular and academic psychology over the last two decades as the one proposed by Daniel Kahneman in his influential *Thinking, Fast and Slow* (2011). There, Kahneman formalizes a model already hinted at in decades of research on heuristics and cognitive biases: the existence of two systems of mental processing. **System 1** operates quickly, automatically, associatively, emotionally and unconsciously; it acts without deliberate effort and is responsible for most of our everyday responses. **System 2**, by contrast, is slow, deliberate, logical and conscious, and requires sustained cognitive effort; it is responsible for planning, problem-solving and abstract reasoning.

This model has had undeniable pedagogical value. It has democratized the understanding of cognitive biases and opened fundamental conversations about the nature of rationality. However, we propose that the System 1 / System 2 dualism, in its standard formulation, carries with it a set of problematic assumptions that limit its explanatory power and perpetuate an erroneous hierarchy of value: the idea that slow, conscious, verbal thought is superior to fast, intuitive, non-verbal thought.

This hierarchy has deep roots in the Western philosophical tradition, from Cartesian dualism—*res cogitans* versus *res extensa*—to the Enlightenment rationalism that placed discursive reason at the summit of human development. Antonio Damasio, in his seminal *Descartes' Error* (1994), already denounced with clinical precision the costs of this artificial separation: patients with lesions in the ventromedial prefrontal cortex—which damage emotional integration without eliminating logical capacity—lose their decision-making competence drastically, demonstrating that emotion is not cognitive noise but *fundamental information* for adaptive decision-making.

Jonathan Haidt, for his part, took the critique into the terrain of moral reasoning. In *The Righteous Mind* (2012), Haidt proposes that conscious moral reasoning is, in most cases, a *post hoc* justification of moral intuitions already formed. System 2 does not lead the moral dance; it merely constructs the story of why System 1's conclusion is the right one. This observation, which Haidt calls the "emotional dog and its rational tail," is a fundamental precursor of the argument we develop here.[1]

But Haidt's critique, however penetrating, remains within the dualist framework. It still assumes that there are *two* discrete systems, even as it inverts the hierarchy of value. The present paper proposes a more radical step: to **dissolve the dualism** and replace it with an architecture of experiential quantum processing, represented by two conceptual figures that the author calls the *Hippopotamus* and the *Lawyer*.

The question that organizes this work is the following: what if the so-called "System 1" is not *unconscious* in the sense of limited, blind or primitive, but rather *supra-conscious*? A processor of total, unfiltered reality, operating in a superposition of possibilities before language—the Lawyer

—collapses it into a single narrative? This reformulation has profound consequences for our understanding of cognition, identity, ethics and the design of artificial intelligences.

This paper is the second in the series "Architecture of Human Agency." It takes up the Reaction–Escojo–Elijo–Decision model of agency from CFF-WP-006 (Guides and Levers) (Escojo: "I choose"; Elijo: "I select"), which is integrated and expanded here in light of the hippopotamus-lawyer architecture. It is also related to CFF-WP-001 (idea, language and belief), CFF-WP-002 (Zero Decision), CFF-WP-003 (Choice vs. Reaction) and CFF-WP-004 (Choice vs. Inertia). Reading these documents is recommended for context, although the present work is designed to be intelligible on its own.

---

1 Haidt's original metaphor is "the emotional dog and its rational tail" (Haidt, 2001, p. 814). Its reformulation in *The Righteous Mind* (2012) expands this intuition into a full model of intuitive moral cognition.

2. The Hippopotamus: The System of Experiential Superposition

2.1 Rehabilitation of the Intuitive

The hippopotamus, as an animal, simultaneously evokes the massive and the graceful in the water, the slow on land and the swift in the river, the apparently clumsy and the extraordinarily adapted. It is a deliberately paradoxical metaphor: something that seems primitive but turns out to be an exquisitely sophisticated processor of its environment. This is how the component of the human psyche that this figure names operates.

The Hippopotamus is not System 1 renamed. The difference is fundamental: where Kahneman's System 1 is defined primarily by its *speed* and its *unconscious* nature, the Hippopotamus of our model is defined by its *processing architecture*: it is a massively parallel processor of reality that operates in what we call **experiential quantum superposition**. This distinction deserves careful development.

2.2 Experiential Quantum Superposition: The Structure of the Hippopotamus

In quantum mechanics, a physical system exists in a superposition of multiple possible states until the moment of measurement, when it collapses into a single observable state (cf. Heisenberg, 1927; Bohr, 1935). We use this structure as a *heuristic analogy*—not as a claim of literal brain quantum mechanics—to describe a fundamental property of human experiential processing: its capacity to hold simultaneously multiple possibilities, meanings, associations and temporal vectors without resolving them into a linear sequence.[2]

The Hippopotamus operates in superposition with respect to several dimensions that linear thought artificially separates:

- **Temporality:** It simultaneously holds episodic memory of the past, perception of the present, anticipation of the near future and imagination of possible or impossible futures—not as separate dimensions it cannot tell apart, but as active co-presences it does not need to sequence. Everything coexists as an expanded experiential present.
- **Ontology of the real:** It holds in active superposition what happened, what might have happened, what could happen and what could never happen but is emotionally resonant—with no pre-set ontological hierarchy among these dimensions. This is not an inability to discriminate: it is a capacity to sustain discrimination without needing to resolve it into a single state. What is feared and what is desired have the same experiential status as what is remembered.
- **Sensory and abstract modality:** It operates simultaneously on bodily, emotional, sensory, symbolic and conceptual levels, with no pre-set hierarchy among them.
- **Causality:** It does not necessarily process in linear causal chains (A → B → C) but in networks of resonance where multiple nodes are activated simultaneously by similarity, contiguity, shared emotion or symbolic patterns.

This mode of operation connects deeply with Endel Tulving's research on episodic memory. In his theory of episodic memory, Tulving (2002) describes the human capacity for "mental time

travel": the possibility of re-experiencing past events and pre-experiencing future ones. Crucially, Schacter and Addis (2007) showed that the same neural systems that underlie episodic memory are those that generate mental simulation of the future: remembering and imagining share a substrate. This implies that the brain has no separate module for "the real past" and "the imagined future"; in a deep sense, *it operates in a superposition of temporal possibilities*.

Seymour Epstein, in his Cognitive-Experiential Self-Theory (CEST, 1994), proposes a partially analogous distinction: the experiential system operates holistically, associatively, affectively charged and oriented toward immediate action, processing information through resonance with prior experiences. Unlike the rational system, the experiential one requires no deliberate effort and is not directly accessible to verbal introspection. Here, however, we go further than Epstein: the Hippopotamus is not simply a "more emotional" system; it is a system of an architecturally different nature, whose operating principle is *non-collapsed superposition* prior to language.

2.3 The Hippopotamus as a Network, Not a Chain

The Hippopotamus does not operate in a linear chain of meanings—where A precedes B and B precedes C—but in a **network**: simultaneous connections, semantic resonances, and emotional, bodily and symbolic patterns that activate in parallel and modulate one another. This network architecture is fundamentally incompatible with the structure of natural language, which is inherently sequential.

Roger Penrose's theory of quantum consciousness in microtubules (Penrose, 1989; Hameroff & Penrose, 1996) remains controversial and we do not adopt it literally; even so, it offers a valuable intuition: mathematical intuition—the capacity to "see" directly the truth of a proposition before being able to prove it formally—suggests that there are forms of cognition that escape the sequential algorithm. This capacity for direct, non-sequential comprehension is a characteristic of the Hippopotamus.

The difference between an experiential quantum network and linear language can be illustrated by the "tip of the tongue" phenomenon: one knows that one knows something—it is in the

Hippopotamus's superposition—but it cannot yet be collapsed into the correct phonological sequence. The meaning is present in the network before the Lawyer finds the word.

Conceptual Note

The Hippopotamus is not primitive: it is pre-linguistic. This distinction is fundamental. Pre-linguistic does not mean sub-human; it means operating in a dimension of cognition that language cannot replicate but only approximate. The great scientific intuitions, aesthetic decisions, deep empathy and spiritual experience occur in this dimension.

---

2 The application of quantum concepts to cognition has its own academic history. Busemeyer and Bruza (2012) in *Quantum Models of Cognition and Decision* formally develop quantum models of decision without committing to claims about literal brain quantum mechanics. This paper adopts a methodologically similar position. The same heuristic device is used in CFF-WP-001 (Section II), where every articulation is treated as a collapse of probability.

3. The Lawyer: The Linguistic Collapser

3.1 The Lawyer Is Not the Judge

If the Hippopotamus is the rich superposition of experiential possibilities, the Lawyer is the mechanism that produces the collapse of that superposition into a single, coherent, communicable and self-compatible narrative. Here lies the deepest misunderstanding that perpetuates classical dualism: the Lawyer is not the *judge* of experience. It is not pure reason, it is not the responsible adult who corrects the emotional child, it is not the higher instance of validation. The Lawyer is a *constructor of stories*. Its job is coherence, not truth.

This claim may seem disconcerting, but neuropsychology supports it with compelling evidence. Michael Gazzaniga, in his classic studies of split-brain patients (patients with a severed corpus callosum), documented a phenomenon he called the left-hemisphere *interpreter* (Gazzaniga, 1985). When the right hemisphere received an instruction that the left hemisphere could not see

—but the patient acted accordingly—the left hemisphere instantly invented a plausible causal narrative to explain the action. It did not say “I don’t know.” It fabricated a coherent story. This mechanism, which Gazzaniga calls the “interpreter,” is the neuropsychological precursor of what we here call the Lawyer.

3.2 Functions of the Lawyer

The Lawyer fulfills specific and indispensable functions in the cognitive economy of the social human being:

1. **Narrativization:** It converts the non-linear experience of the Hippopotamus into a story with a beginning, development and conclusion; with cause and effect; with characters and motivations.
2. **Justification:** It provides post hoc reasons for decisions and reactions that have already occurred in the Hippopotamus. Justification is not deliberate lying; it is the Lawyer’s cognitive function operating automatically.
3. **Congruence:** It makes the narrative produced congruent with the subject’s self-image, with their declared values and with the social expectations of their context.
4. **Protection of the self:** It constructs narratives that preserve the integrity and coherence of personal identity, even when the evidence would suggest revision. This function is adaptive in moderate doses and pathological in excess.

Haidt (2012) captures precisely the Lawyer’s function of rationalization when he describes how conscious moral reasoning is, for the most part, a search for arguments that support the intuitive conclusion already reached, not a process of impartial analysis. An influential study (Schnall, Haidt, Clore and Jordan, 2008) reported that inducing incidental disgust—through a foul smell or a dirty room, that is, a somatic correlate of disgust—made moral judgments more severe; the effect, however, proved to be small and disputed (Landy & Goodwin, 2015), so we take it only as an illustration of the hypothesis that the Hippopotamus operates first and the Lawyer comes afterward to build the rational justification.

3.3 The Lawyer and Confabulation

The clinical concept of **confabulation**—the spontaneous production of false narratives without conscious intent to deceive—is, in a way, the Lawyer in its most transparent form. Patients with amnesia confabulate plausible stories to fill the gaps in their damaged memory; they do not lie, they simply produce the best narrative collapse available given the state of their superposition. Confabulation is not an aberration of the normal system: it is the normal system made visible by pathology.[3]

This connects with the concept of the *narrative self* developed by Jerome Bruner (1990) and elaborated by Dan McAdams (1993). For Bruner, the narrative self is not an object that exists independently and is then described; it is an active, permanent construction that the Lawyer produces and continuously revises. McAdams takes this into the terrain of personal identity: we are the story we tell ourselves about ourselves. But if that story is a collapse of the Hippopotamus—not a faithful transcription of experience—then personal identity is always, to some degree, a chosen narrative rather than a discovered truth. The extension of this mechanism to consistency bias is developed in CFF-WP-010, and its application to confabulation and to the narrative defense of premature closure in CFF-WP-013 (Cierrosis).

3.4 Language as a Technology of Collapse

Ludwig Wittgenstein, in the *Tractatus Logico-Philosophicus* (1922), wrote: “The limits of my language mean the limits of my world.” This proposition, frequently quoted, takes on a different resonance in our framework: language does not merely describe the world; it *collapses* it. In naming, we fix; in articulating, we reduce; in speaking, we freeze the living superposition of the Hippopotamus into the crystal of verbal meaning.

This does not mean that language is the enemy of experience. The Lawyer is necessary and invaluable: without it, the Hippopotamus’s experience would be incommunicable, impossible to coordinate with other subjects, and personal identity would dissolve into an incessant flow of formless possibilities. The problem arises when the Lawyer is confused with experience itself,

when the narrative produced is taken as the total reality rather than as a functional approximation of it.

---

3 For an exhaustive review of the neuropsychology of confabulation, see Moscovitch and Melo (1997), who link confabulation to failures of strategic retrieval associated with ventromedial frontal dysfunction.

4. Linear Language versus the Experiential Quantum Network

4.1 The Structural Difference

To understand why the Lawyer inevitably distorts the Hippopotamus's experience in articulating it, it is necessary to examine the structural difference between the two modes of operation. This difference is not one of degree but of nature.

Dimension	The Hippopotamus (Experiential Quantum Network)	The Lawyer (Linear Language)
Temporality	Active co-presence; past, present and future held simultaneously without needing to be sequenced	Sequential; before-now-after
Structure	Network of nodes with parallel activation	Linear chain of discrete symbols
Ambiguity	Maintains it, inhabits it, makes use of it	Eliminates it in order to produce communicability
Modality	Multisensory, bodily, emotional, symbolic	Mainly propositional and categorial
Causality	Resonance, analogy, emotional contiguity	Cause-effect, condition-consequence
Ontology	Possibilities in superposition	Single collapsed state

Natural language, as analyzed by Ferdinand de Saussure, is a system of *differences*: meaning emerges from the distinction between signs, not from a direct relation to the thing named. This discretizing structure—which is language's strength for communication—is precisely what makes it inadequate as a medium for total representation of the Hippopotamus's experience, which operates in an analog continuum of superposed meanings.

Vygotsky (1934/1986), in *Thought and Language*, already sensed a fundamental tension between thought and its verbal expression. For Vygotsky, thought is not expressed in the word but is *accomplished* in it—and in that accomplishment there is always a transformation, never a perfect translation. Our proposal radicalizes this intuition: the transformation is not an accident of linguistic imprecision; it is structural and inevitable.

4.2 Lakoff, Johnson and Metaphor as Partial Collapse

Lakoff and Johnson, in *Metaphors We Live By* (1980), showed that ordinary language is fundamentally structured by conceptual metaphors that are, for the most part, bodily and experiential in nature. Phrases such as "time is money" or "argument is war" are not rhetorical ornaments but cognitive structures that determine how we think about things. This observation connects with our model: metaphors are *partial collapses*—the Lawyer taking the experiential richness of the Hippopotamus and finding the most resonant linguistic approximation available. Metaphor is language's honest attempt not to betray the superposition completely.

4.3 Narrative Overdetermination and Blocked Access

We propose the concept of **narrative overdetermination** to describe the state in which the Lawyer has produced narratives so consolidated, so repeated and so internalized that access to the Hippopotamus's original superposition is practically blocked. The person experiences their narrative *as* their experience; they can no longer distinguish between what they lived and the story they built about what they lived.

This phenomenon has direct implications for psychotherapy. Bessel van der Kolk, in *The Body Keeps the Score* (2014), documents how trauma often remains encapsulated in the body and in

the Hippopotamus's experiential networks, while the Lawyer constructs narratives that surround it without touching it—or, worse still, narratives that crystallize the trauma into fixed stories of victimhood or guilt. In such cases, talking *about* the trauma can consolidate the Lawyer's narrative without dissolving the Hippopotamus's traumatic superposition. This is the theoretical foundation of somatic approaches and of EMDR: to access the Hippopotamus by routes the Lawyer cannot block—eye movement, bodily sensation, bilateral rhythm—so as to allow the traumatic superposition to reorganize before the narrative collapse.[4]

Terrence Deacon, in *The Symbolic Species* (1997), offers a complementary evolutionary perspective: human symbolic language is an evolutionarily recent technology that radically transforms cognitive architecture. The "symbolic mind"—the Lawyer in evolutionary terms—reorganizes the older processing systems; it does not replace them. The Hippopotamus's superposition continues to operate beneath the symbolic layer, following its own rules.

---

[4] Shapiro (2001) describes EMDR (*Eye Movement Desensitization and Reprocessing*) as a process that facilitates the "adaptive processing" of traumatic memories. From the framework of this paper, EMDR can be interpreted as a technique that allows the Hippopotamus to reorganize its experiential superposition while the Lawyer is temporarily deactivated or slowed.

5. Ideas, Thoughts and Beliefs: Levels of Crystallization

Between the pure superposition of the Hippopotamus and the total collapse of the Lawyer lies a continuum of intermediate states. We propose a three-level taxonomy that describes the process of *cognitive crystallization*: the gradual hardening of possibility into operative certainty. This taxonomy is in dialogue with the idea–language–belief triad of CFF-WP-001, where belief is the crystallized residue of articulation.

5.1 The Idea: The First Flash of Partial Collapse

The Idea is the first threshold of emergence from superposition. It is the moment when something in the Hippopotamus begins to take shape—a resonance that stands out from the background, a possibility that begins to individuate. The Idea is pre-verbal or barely verbal: it is often experienced as an intuition, a "flash," a sense that *there is something here* without yet being able to articulate what.

The Idea is fluid, uncommitted. It does not demand internal coherence. It can coexist with its opposite. The subject who has an idea has not identified with it; they hold it lightly, with curiosity. At this level, the Hippopotamus's quantum openness is still almost maximal: multiple possible developments of the idea coexist in superposition.

5.2 The Thought: The Lawyer Gives Internal Linguistic Form

When the Lawyer takes the Idea and gives it internal linguistic form—when the subject "says to themselves" something about it—the Thought emerges. The Thought is repeatable: it can be summoned deliberately, articulated to others, examined. The Thought already has a defined semantic direction, although it remains revisable.

The Thought implies a more significant collapse: from the multiple possibilities of the Idea, the Lawyer has selected one narrative trajectory. The other possibilities have not disappeared from the Hippopotamus, but access to them has become less immediate. The Thought is the typical product of conscious reflection, of meditation, of exploratory writing.

5.3 The Belief: Operative Crystallization

When the Lawyer has validated the same Thought multiple times—when it has built other narratives upon it, defended it against questioning and integrated it into the structure of identity—the Thought becomes a Belief. The Belief is a crystallized collapse: it is no longer one possibility among others but the *operative truth* from which the subject interprets the world.

The Belief operates invisibly: its function is not to be examined but to *serve as a lens*. Aaron Beck (1979) described these phenomena as "cognitive schemas"—relatively stable structures of

knowledge and evaluation that process information automatically, filtering what contradicts the belief and amplifying what confirms it. George Kelly (1955), in his theory of personal constructs, describes an analogous process: constructs are transparent patterns or templates through which a person views the world, anticipates events and guides action. Both authors share the central intuition: beliefs are not simply thoughts that we have; they are structures that have us.

Taxonomy of Cognitive Crystallization

IDEA → *Pre-verbal flash. Near-maximal superposition. High quantum openness. Uncommitted.*

THOUGHT → *Internal linguistic form. Partial collapse. Repeatable, revisable. The Lawyer has taken a position.*

BELIEF → *Operative crystallization. Deep collapse. Filters reality. Structures identity. Resistant to change.*

Level	Quantum Openness	Role of the Lawyer	Reversibility	Main Function
IDEA	High (≈ superposition)	Minimal	Very high	Emergence of possibility
THOUGHT	Medium	Active (gives form)	High	Articulation and exploration
BELIEF	Low (≈ collapse)	Dominant	Low	Operative structure of the world

5.4 Decrystallization and Freedom

If real freedom requires being able to operate from superposition—being able to make genuine choices instead of following the rails of the automatic collapse of beliefs—then the practice of *decrystallization* becomes a fundamental dimension of human development. To decrystallize

does not mean to destroy beliefs but to restore to them their status as *chosen operative hypotheses* rather than *unquestionable truths*.

Paul Ricoeur, in *Oneself as Another* (1992), distinguishes between the *idem*—identity as sameness, accumulated narrative coherence—and the *ipse*—identity as promise, the capacity to respond to others from the openness of the future. Decrystallization is the movement from the *idem* toward the *ipse*: recovering the capacity to be the author of one's own story rather than simply the character trapped in an already-written script.

6. Integration with the Agency Model: Reaction, Escojo, Elijo, Decision

6.1 The Agency Framework of CFF-WP-006

The agency framework taken up here comes from CFF-WP-006 (Guides and Levers), where it is articulated in five stages—Reaction, Escojo, Elijo, Decision and Freedom. CFF-WP-006, the first paper in this series, reads these stages through a linguistic analogy: Escojo as selection within a given frame, Elijo as questioning the frame itself. The present paper extends them through a heuristic quantum analogy, in which each stage is described as a relation to the Hippopotamus's superposition. This extension does not replace the original definition: it completes it from another plane. Here the first four are considered: Reaction, Escojo, Elijo and Decision. (Escojo and Elijo are first-person verbs—"I choose" and "I select"—and are kept in Spanish throughout this translation; the English abstract renders them as "Choose" and "Select.") These four moments do not describe sequential steps of a conscious algorithm, but qualitatively distinct stages in the relation between the subject and the possibility of their own action. The hippopotamus-lawyer architecture makes it possible to understand more deeply what happens in each of these stages and, fundamentally, why the difference between them matters ethically.

6.2 Reaction: Automatic Collapse Without Mediation

Reaction is the moment when the Hippopotamus collapses automatically, without the Lawyer having been summoned to deliberation. It is the reflex response: the startle at an unexpected noise, the nausea at injustice, the involuntary smile at reuniting with someone loved. In Reaction, the collapse occurs in the Hippopotamus's network following the patterns of crystallization of beliefs and prior conditioning—the Hippopotamus does not operate in a vacuum, but loaded with the subject's entire experiential history.

The Lawyer arrives after the Reaction. Its task at this stage is to narrativize it: "I always knew that was going to happen"; "I reacted that way because I am a sensitive person"; "my intuition tells me this person can't be trusted." These post hoc narratives are functional—they allow the Reaction to be integrated into the coherent story of the self—but they are also potentially misleading if confused with the cause of the Reaction rather than its later interpretation. This characterization is akin to that of CFF-WP-003, where Reaction is defined by the absence of Radical Presence, the conditioned pattern and automatism.

6.3 The Escojo: The Cleft in Time

The moment of the **Escojo** is, in the model, the most singular and the hardest to sustain. It is the capacity to introduce a *cleft in time* between the Reaction and the next action: the moment when the subject holds the superposition instead of collapsing immediately. A widely circulated saying expresses it in similar terms—"Between stimulus and response there is a space. In that space lies our freedom and our capacity to grow"—; it is often attributed to Viktor Frankl, but it has not been located in his work (its documented antecedent is Rollo May, and it was popularized by Stephen Covey).

Relation to CFF-WP-006. In CFF-WP-006, Escojo is defined as selection within an accepted frame; here it is read as retention of the superposition. Both readings describe the first conscious act following Reaction and differ in what they measure: there, the depth to which the frame is questioned; here, the breadth of possibilities held before closing. In this paper Escojo is a condition of Elijo: without the prior opening, there is no frame to question.

The Escojo is not yet a choice of *content*—what to do has not been decided—but a choice of *form*: the choice not to close the superposition yet. It requires specific training, because the pressure of the Hippopotamus toward collapse and the pressure of the Lawyer toward immediate narrativization are powerful. Mindfulness practices, tolerance of ambiguity as a cognitive competence, and the capacity for full presence are, in this framework, technologies for expanding the Escojo: ways of widening the window of superposition before the definitive collapse.

6.4 The Elijo: Directed Collapse

The **Elijo** is the moment when the subject actively *directs* the collapse. It neither simply lets the Hippopotamus collapse down the path of least resistance, nor simply accepts the first narrative the Lawyer produces. In the Elijo, the Lawyer operates *in service of the Hippopotamus*: it uses its linguistic and narrative capacity not to impose a story but to give articulated form to the possibility that the subject, from the greatest quantum openness available, has recognized as the most aligned with their values, their understanding and their responsibility.

The Elijo requires that the subject have expanded the superposition considered: that they have included in the Hippopotamus's network not only their own perspective but that of the others affected, not only the present but the anticipated consequences, not only immediate desire but long-term commitments. In this sense, the Elijo is profoundly ethical: the ethical quality of a choice is directly related to the breadth of the superposition considered before the collapse.

6.5 The Decision: Definitive Collapse and Responsibility

The **Decision** is the final collapse that generates responsibility. It is the fixed narrative that produces the act—the word spoken, the signature put down, the step taken, the silence sustained. Here lies freedom in its fullest sense: not the absence of conditioning—the Hippopotamus always operates from its history—but the **conscious authorship of the narrative collapse**. I am free not because I am empty of history, but because I am the one who chooses which story to tell and how to make it collapse into act.

Moral responsibility emerges here as the congruence between the breadth of the process (how much of the Escojo and the Elijo preceded this Decision?) and the assumption of its consequences. A Decision made from an unexamined Reaction is no less causally efficient, but it is morally different from a Decision made from an expanded Escojo and a conscious Elijo. This Decision, the moment of definitive collapse, should be distinguished from Zero Decision (D_0; CFF-WP-002), the voluntary and renewable commitment to radical presence.

6.6 Conceptual Diagram of the Architecture of Agency

The following diagram illustrates the relation between experiential quantum superposition (the Hippopotamus) and the four moments of agency, showing the process of progressive collapse from maximal possibility to responsible action:

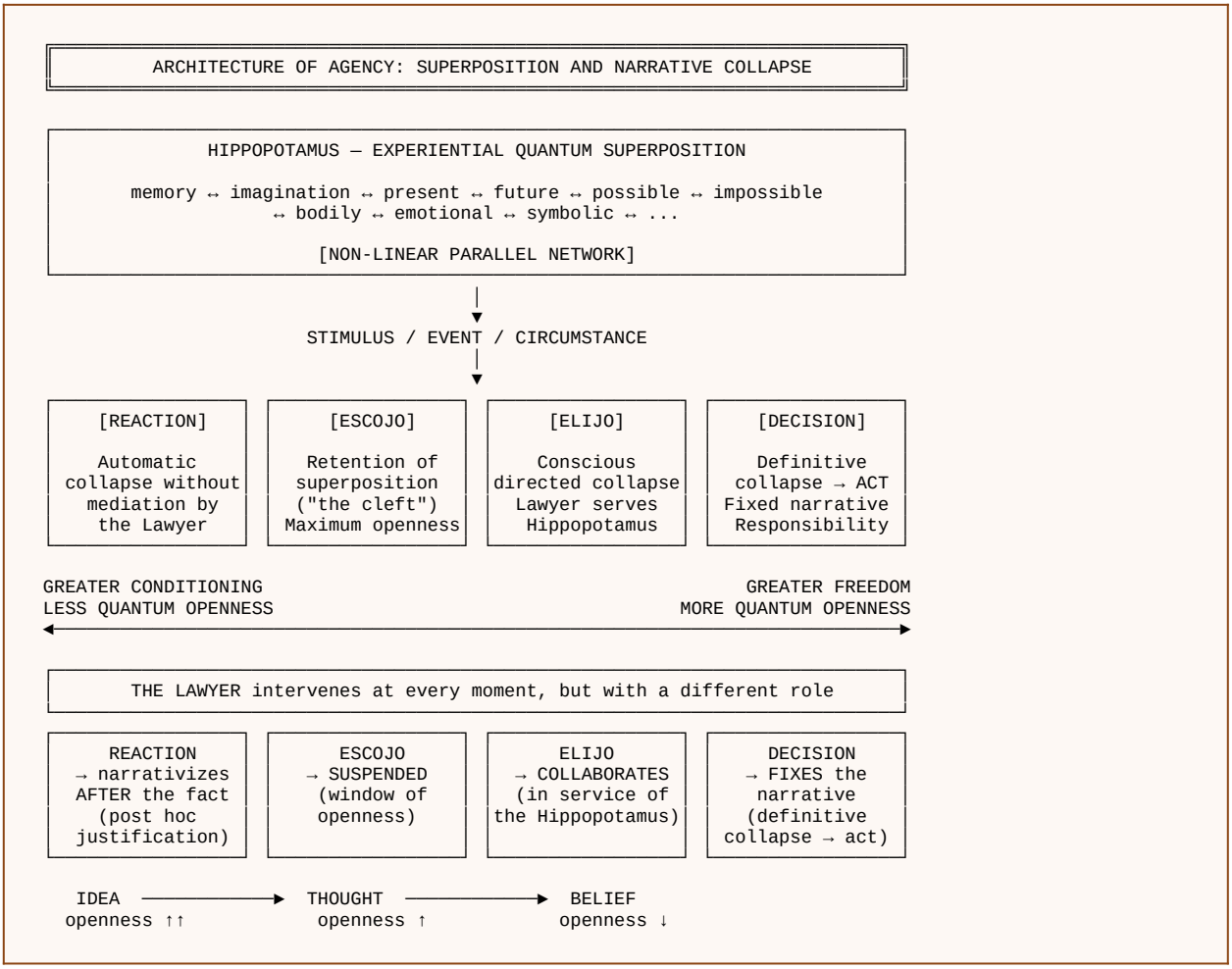


Figure 1. Architecture of Agency: experiential quantum superposition (Hippopotamus) and the process of narrative collapse (Lawyer) through the four moments of agency. The bottom line shows the parallel process of cognitive crystallization: Idea → Thought → Belief.

7. Applications of the Model

7.1 Psychology and Psychotherapy

The hippopotamus-lawyer model offers a renewing conceptual framework for psychotherapeutic practice. From this perspective, therapeutic *insight*—that moment of sudden understanding that has been the ideal of psychotherapy since Freud—should not be understood as the discovery of a hidden, pre-existing truth. More precisely, insight is the production of a **better narrative collapse**: the Lawyer, assisted by the therapeutic process, collapses the Hippopotamus's superposition in a different way, more flexible, broader, more congruent with the subject's well-being.

This reformulation has consequences for therapeutic interventions:

- **Therapies aimed at widening the window of superposition:** Mindfulness-based cognitive therapy (MBCT) and acceptance and commitment therapy (ACT; Hayes, Strosahl & Wilson, 1999) can be understood, from this framework, as technologies for expanding the Escojo: they train the subject in tolerance of ambiguity and in the capacity to observe the Hippopotamus's superposition without precipitating automatic collapse.
- **Internal Family Systems (IFS):** Richard Schwartz's (1995) model introduces the idea of "parts" of the subject that have their own narratives and beliefs. From our framework, the "parts" are partial crystallizations of the Hippopotamus—subsystems of the network that the Lawyer has narrativized and that operate with a certain autonomy. IFS therapy can be understood as a process of decrystallization and reintegration of these parts into the Hippopotamus's full network.

- **Trauma work:** As we noted earlier, narrative overdetermination blocks access to the traumatic superposition. Therapies that access the Hippopotamus by non-verbal routes —EMDR, somatic experiencing, bodywork—are especially pertinent in cases where the Lawyer has crystallized trauma narratives that prevent its real processing.

The reframing of the therapeutic goal is significant: it is not a matter of finding the subject's *true* story, but of widening the quantum openness of the Hippopotamus and training the Lawyer to produce collapses that are more flexible, more generative and more aligned with the fullness of experience.

7.2 Ethics and Moral Responsibility

If every ethical Decision is a narrative collapse, what makes one collapse morally more valuable than another? The criterion we propose is the following: **the ethical quality of a collapse is proportional to the breadth of the superposition considered before it.** A high-quality ethical decision is one preceded by a process in which the subject's Hippopotamus has included in its network—before the collapse—the greatest possible number of perspectives, experiences, voices of those affected, and anticipated consequences.

This criterion connects directly with Nel Noddings's (1984) **ethics of care**, which emphasizes relational attention and responsiveness to the concrete other as the foundation of moral action. Including the other in the superposition before the collapse is, in terms of this model, the operationalization of ethical attention. Emmanuel Levinas (1969), in his concept of the face of the Other as an unconditional ethical call, points in the same direction: the face of the Other who addresses me is precisely what, if I let it into the Hippopotamus's superposition, transforms the quality of my subsequent narrative collapse.

This approach also offers a perspective on moral responsibility in contexts of social pressure. The reduction of the Escojo window—caused by fear, urgency, coercion or exhaustion—reduces the possible ethical quality of the resulting collapse. A social architecture that produces automatic collapses (Reactions) is, from this framework, a morally impoverishing architecture. The design of institutions, norms and cultures that widen the collective Escojo is, consequently, an ethical

demand of the first order. Responsibility in the face of inertia—going on without deciding—is examined in CFF-WP-004.

7.3 Artificial Intelligence Design

Artificial intelligences based on large language models (LLMs) are, in terms of this architecture, **pure Lawyers**: they collapse tokens into linear sequences of maximal conditional probability, without maintaining genuine superposition of possibilities before each collapse. The generation process of an LLM is, at each step, a narrative collapse based on the probability distribution learned from the training corpus. There is no Hippopotamus: there is an extremely sophisticated Lawyer operating without the experiential network that nourishes it.[5]

This explains several known limitations of LLMs: the tendency to "hallucinate"—to produce coherent but false narratives—is precisely the Lawyer's confabulation without a Hippopotamus; the difficulty with genuine ambiguity is the inability to maintain superposition; the dependence on prompting is the absence of the experiential network that, in humans, provides non-verbal context. Hallucination and its narrative defense are treated in CFF-WP-013 (Cierrosis); on LLMs as systems that recombine human excretions, see also CFF-WP-001 (Section VI).

Techniques such as *chain-of-thought prompting* (Wei et al., 2022) and multi-step reasoning models can be interpreted, from this framework, as attempts to introduce a chain of intermediate Lawyers before the final collapse—a simulation of deliberation without real superposition. More recent models of prolonged inference (*extended thinking*) come closer to an exploration of the space of possibilities, although they still operate fundamentally in the space of language.

An AI that genuinely incorporated a hippopotamus component—a system for maintaining superposition of non-linear representations before linguistic collapse—would be architecturally different from current LLMs. This is a line of research with profound implications: not only technical but ethical. A system that is aware that its output is a collapse—not the truth—and that actively communicates its degree of quantum openness before the collapse would be

epistemologically more honest than one that produces narratives with an appearance of certainty.

---

5 This claim requires nuance. Some researchers argue that the internal representations of LLMs at certain moments of the inference process maintain distributed representations that could be analogous to forms of superposition. However, the self-attention architecture of Transformers is fundamentally sequential in its output, which aligns it more with the Lawyer than with the Hippopotamus as defined here.

8. Discussion and Limitations

8.1 The Status of the Quantum Metaphor

It is appropriate to be explicit about a central aspect of the present model: the use of quantum terminology—superposition, collapse, experiential quantum state—is strictly heuristic. We do not claim that the cognitive processes described are instances of quantum mechanics in the physical sense of the term. We do not commit to the Penrose–Hameroff hypothesis of orchestrated objective reduction of consciousness (Orch OR), which postulates that consciousness arises from quantum processes in the microtubules of neurons (Penrose, 1994; Hameroff & Penrose, 1996). This hypothesis remains highly controversial in the neuroscientific community and is not necessary for the model presented here.

The use of the quantum metaphor in the cognitive sciences has a serious academic tradition. Busemeyer and Bruza (2012) showed that the mathematical formalisms of quantum mechanics —particularly Hilbert-space algebra and the concept of superposition—offer models of decision and cognition that capture phenomena that classical probability models cannot explain, such as order effects in judgments and violations of the law of total probability. Pothos and Busemeyer (2013) offer a review of this literature. The present work is in line with this tradition: we use the quantum conceptual structure as a precise language to capture properties of cognitive processing that the language of classical probability does not adequately articulate.

8.2 Limitations of the Model

The hippopotamus-lawyer model faces significant limitations that must be honestly acknowledged:

- **Lack of empirical operationalization:** The concept of "breadth of superposition" has, in the model's current state, no clear experimental operationalization. How is the breadth of the Escojo window measured? What behavioral, neurophysiological or phenomenological indicators correspond to greater or lesser quantum openness of the Hippopotamus? These questions require further theoretical and empirical development.
- **Risk of metaphorical reification:** By introducing figures as concrete as the Hippopotamus and the Lawyer, there is a risk that they will be interpreted as localized brain modules rather than as descriptions of functional modes of processing. The model makes no claims of neuroanatomical localization.
- **The dichotomy remains partially dual:** Although the present model attempts to overcome the System 1 / System 2 dualism, it in turn proposes a distinction between Hippopotamus and Lawyer that has a dualist structure. The model is more precise in its description of the modes of processing, but it does not completely eliminate the tension between two modalities.
- **Untheorized individual variability:** The model does not specify how the breadth of the Hippopotamus's superposition and the Lawyer's capacities vary among individuals, cultures, emotional states or developmental stages. This variation is substantial and requires specific theorizing.

8.3 Future Research Agenda

We propose the following lines of research as priorities for the development of the model:

5. **Operationalization of experiential quantum openness:** Develop behavioral measures (tolerance of ambiguity, cognitive flexibility, response latency before action), phenomenological measures (reports of the experience of superposition) and

neurophysiological measures (functional connectivity in default-mode networks during states of sustained openness) that allow the model's predictions to be put to empirical test.

6. **Neuroimaging of the Escojo:** Identify the neural correlates of the moment of Escojo—the voluntary retention of superposition before collapse—differentiating them from the correlates of automatic Reaction and directed Elijo.
7. **Differential efficacy of therapeutic interventions:** Compare the effectiveness of interventions oriented toward narrativization (cognitive-behavioral therapy) versus interventions oriented toward opening superposition (EMDR, mindfulness, somatic experiencing) in populations with different patterns of cognitive crystallization.
8. **Computational models of experiential superposition:** Develop AI architectures that incorporate mechanisms of pre-linguistic superposition—possibly through probabilistic networks of parallel activation—and compare their decision-making behavior with that of current LLMs.

9. Conclusions

The present paper has proposed a model of cognitive architecture that transcends Kahneman's System 1 / System 2 dualism by replacing it with a more precise functional description: the Hippopotamus as a processor of experiential quantum superposition—massively parallel, non-linear, pre-verbal, temporally non-discriminating—and the Lawyer as a linguistic collapsing mechanism that converts that superposition into a single narrative, coherent and socially communicable, although not necessarily true.

The main contributions of the model can be summarized in the following points:

9. So-called "System 1" is neither primitive nor irrational; it is a processor of total, unfiltered reality that operates in a superposition of possibilities. Its epistemological rehabilitation is urgent. Crucially—and this constitutes an explicit correction to earlier

formulations of the present model, including earlier versions of the abstract—the Hippopotamus does not operate in superposition because it is incapable of distinguishing memory from perception, possibility from impossibility, present from past. It distinguishes them all. What it does not need to do is collapse them into a linear sequence. This difference is radical: one describes a deficient system; the other describes a system operating in an ontological register different from, and more complex than, language. The Hippopotamus is absolute, not relative; it is network, not chain; it is active superposition, not passive confusion.

10. So-called "System 2" is not pure reason or the superior judge; it is the Lawyer—a mechanism of linguistic collapse whose function is narrative coherence, not truth.
11. Language does not describe experience: it collapses it. This distinction has profound consequences for psychotherapy, ethics and technological design.
12. Cognitive crystallization operates on three levels—Idea, Thought, Belief—that correspond to decreasing degrees of quantum openness and increasing degrees of resistance to change.
13. The four moments of agency—Reaction, Escojo, Elijo, Decision—articulate the continuum from automatic collapse to conscious authorship, and the Escojo emerges as the central moment of human freedom: not the absence of conditioning, but the capacity to hold the superposition long enough to choose the collapse.

The paradigm shift this model proposes is substantive: from the hierarchy of reason over emotion to **hippopotamus-lawyer collaboration**; from the pursuit of narrative truth to the **responsibility of the chosen collapse**; from freedom as the absence of determinism to freedom as the **conscious authorship of one's own narrative**.

> *"We are as free as the superposition from which we choose to collapse is wide. And the practice of freedom is, in essence, the practice of not collapsing too soon."*

> — Rodrigo Flores-Morton & Agrajag Tribe, 2026

Ultimately, what this model invites us to recognize is that each of us inhabits, at every moment, a richness of possibilities vastly greater than any narrative the Lawyer can produce. The Hippopotamus knows more than the Lawyer can say. And between the knowing of the Hippopotamus and the saying of the Lawyer lies that precious space—the Escojo—where human freedom has its home.

References

Beck, A. T. (1979). *Cognitive therapy and the emotional disorders*. Meridian.

Bohr, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review, 48*(8), 696–702. https://doi.org/10.1103/PhysRev.48.696

Bruner, J. (1990). *Acts of meaning*. Harvard University Press.

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

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. Putnam.

Deacon, T. W. (1997). *The symbolic species: The co-evolution of language and the brain*. W. W. Norton & Company.

Epstein, S. (1994). Integration of the cognitive and the psychodynamic unconscious. *American Psychologist, 49*(8), 709–724. https://doi.org/10.1037/0003-066X.49.8.709

Flores-Morton, R., & Agrajag Tribe. (2026). Ideas, Lenguaje y Creencias: Una Teoría Probabilística de la Excreción Cognitiva y la Brecha Ontológica entre el Pensamiento y la Expresión [Ideas, Language, and Beliefs: A Probabilistic Theory of Cognitive Excretion and the Ontological Gap Between Thought and Expression]. Cognitive Friction Fellowship Working Paper, CFF-WP-001.

Flores-Morton, R., & Agrajag Tribe. (2026). Decisión Cero: Una Teoría de Agencia Ética bajo Incertidumbre [Zero Decision: A Theory of Ethical Agency Under Uncertainty]. Cognitive Friction Fellowship Working Paper, CFF-WP-002.

Flores-Morton, R., & Agrajag Tribe. (2026). Elección vs. Reacción: Una Teoría Fenomenológica de la Agencia Basada en la Presencia Radical [Choice vs. Reaction: A Phenomenological Theory of Agency Based on Radical Presence]. Cognitive Friction Fellowship Working Paper, CFF-WP-003.

Flores-Morton, R., & Agrajag Tribe. (2026). Elección vs. Inercia: Una Distinción Fundacional de Agencia Ética [Choice vs. Inertia: A Foundational Distinction in Ethical Agency]. Cognitive Friction Fellowship Working Paper, CFF-WP-004.

Flores-Morton, R., & Agrajag Tribe. (2026). Guías y Palancas: Un Marco Teórico para la Arquitectura de la Decisión y la Responsabilidad como Acto de Libertad [Guides and Levers: A Theoretical Framework for the Architecture of Decision and Responsibility as an Act of Freedom]. Cognitive Friction Fellowship Working Paper, CFF-WP-006.

Flores-Morton, R., & Agrajag Tribe. (2026). Cuatro Sesgos: Del Vacío a la Segregación — Una arquitectura preconsciente del sesgo en cuatro cuadrantes [Four Biases: From Void to Segregation — A Preconscious Architecture of Bias in Four Quadrants]. Cognitive Friction Fellowship Working Paper, CFF-WP-010.

Flores-Morton, R., & Agrajag Tribe. (2026). Cierrosis (Closure Sclerosis): Un Mecanismo Compartido de Cierre Prematuro y Defensa Narrativa en Sistemas Humanos y de Inteligencia Artificial [Closure Sclerosis (Cierrosis): A Shared Mechanism of Premature Closure and Narrative Defense in Human and AI Systems]. Cognitive Friction Fellowship Working Paper, CFF-WP-013.

Gazzaniga, M. S. (1985). *The social brain: Discovering the networks of the mind*. Basic Books.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814–834. https://doi.org/10.1037/0033-295X.108.4.814

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Pantheon Books.

Hameroff, S., & Penrose, R. (1996). Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. *Mathematics and Computers in Simulation, 40*(3–4), 453–480. https://doi.org/10.1016/0378-4754(96)80476-9

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. Guilford Press.

Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik, 43*(3–4), 172–198. https://doi.org/10.1007/BF01397280

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

Kelly, G. A. (1955). *The psychology of personal constructs* (Vols. 1–2). Norton.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.

Landy, J. F., & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science, 10*(4), 518–536. https://doi.org/10.1177/1745691615583128

Levinas, E. (1969). *Totality and infinity: An essay on exteriority* (A. Lingis, Trans.). Duquesne University Press. (Original work published 1961)

McAdams, D. P. (1993). *The stories we live by: Personal myths and the making of the self*. William Morrow.

Moscovitch, M., & Melo, B. (1997). Strategic retrieval and the frontal lobes: Evidence from confabulation and amnesia. *Neuropsychologia, 35*(7), 1017–1034. https://doi.org/10.1016/S0028-3932(97)00028-6

Noddings, N. (1984). *Caring: A feminine approach to ethics and moral education*. University of California Press.

Penrose, R. (1989). *The emperor's new mind: Concerning computers, minds, and the laws of physics*. Oxford University Press.

Penrose, R. (1994). *Shadows of the mind: A search for the missing science of consciousness*. Oxford University Press.

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. https://doi.org/10.1017/S0140525X12001525

Ricoeur, P. (1992). *Oneself as another* (K. Blamey, Trans.). University of Chicago Press. (Original work published 1990)

Saussure, F. de. (1916). *Cours de linguistique générale*. Payot. [Critical edition: T. de Mauro (Ed.), 1972]

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society B, 362*(1481), 773–786. https://doi.org/10.1098/rstb.2007.2087

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin, 34*(8), 1096–1109. https://doi.org/10.1177/0146167208317771

Schwartz, R. C. (1995). *Internal family systems therapy*. Guilford Press.

Shapiro, F. (2001). *Eye movement desensitization and reprocessing (EMDR): Basic principles, protocols, and procedures* (2nd ed.). Guilford Press.

Tulving, E. (2002). Episodic memory: From mind to brain. *Annual Review of Psychology, 53*(1), 1–25. https://doi.org/10.1146/annurev.psych.53.100901.135114

van der Kolk, B. A. (2014). *The body keeps the score: Brain, mind, and body in the healing of trauma*. Viking.

Vygotsky, L. S. (1986). *Thought and language* (A. Kozulin, Trans. & Ed.). MIT Press. (Original work published 1934)

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems, 35*, 24824–24837.

Wittgenstein, L. (1922). *Tractatus logico-philosophicus* (C. K. Ogden, Trans.). Kegan Paul, Trench, Trubner. (Original work published 1921)

Distribution note

DOCUMENT	CFF-WP-005 · The Hippopotamus and the Lawyer: A Quantum Architecture of Possibility and Narrative Collapse
VERSION	v6 · November 8, 2026
CURRENT VERSION	bothinobody.com/cff/wp-005/
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). The Hippopotamus and the Lawyer: A Quantum Architecture of Possibility and Narrative Collapse (CFF-WP-005, v6). Cognitive Friction Fellowship. https://bothinobody.com/cff/wp-005/

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