Concept Paper

A Neurophilosophical Model of Personal and Meta-Reflective Modes of Mind

Kyrylo Somkin *

Posted Date: 5 March 2026

doi: 10.20944/preprints202603.0446.v1

Keywords: neurophilosophy; metacognition; neuroscience; neural integration; consciousness; neurobiology; sociology; psychology

 doi:10.20944/preprints202603.0446.v1

Concept Paper

A Neurophilosophical Model of Personal and Meta-Reflective Modes of Mind

Kyrylo Somkin

Independent Researcher, Czech Republic; somkinkirilo@gmail.com

Abstract

This paper proposes a neurophilosophical conceptual model of human consciousness structured as two functional brain states: the personal mode and the meta-reflective mode. The personal mode is defined as a motivationally and socially embedded configuration of neural processes oriented toward adaptation, identity maintenance, and ego-relevant concerns. The meta-reflective mode is characterized as a functional state in which cognition turns upon itself, enabling abstraction, self-objectification, and existential evaluation. The model does not posit a metaphysical dualism nor strictly separable neural systems. Rather, both modes may recruit overlapping brain regions, including prefrontal structures, while differing in dominant functional orientation and hierarchical organization. The distinction is therefore not anatomical but configurational. It is argued that tensions between these modes may account for different categories of psychological crises: identity-based crises primarily emerging within the personal mode, and existential crises arising from intensified meta-reflective activation. The framework further suggests that the development of civilization reflects the structural coexistence of adaptive engagement and reflective distancing. While empirical validation remains limited, the model aims to provide a structured bridge between phenomenological analysis and contemporary neurocognitive theory.

Keywords: neurophilosophy; metacognition; neuroscience; neural integration; consciousness; neurobiology; sociology; psychology

Introduction

Human experience appears structurally unified yet phenomenologically differentiated. Individuals simultaneously participate in socially conditioned, motivationally driven engagement with the world and in moments of reflective distancing, during which their own beliefs, identity, and even the structure of reality become objects of examination. Although cognitive science has extensively investigated adaptive, affective, and decision-making processes, and philosophy has long analyzed reflective consciousness and self-awareness, the structural relation between these domains remains insufficiently articulated. Contemporary discussions in neurophilosophy and cognitive theory frequently address dual-process models, metacognition, and self-representation. However, many of these frameworks emphasize computational efficiency, information processing, or epistemic evaluation, leaving underdeveloped the phenomenological tension between lived identity and meta-reflective abstraction. In particular, the observable distinction between identity-related crises (social, developmental, affective) and existential crises (ontological, temporal, metaphysical) suggests the possibility of functionally distinct configurations within a unified neurocognitive system.

This article proposes a conceptual neurophilosophical model that distinguishes two functional states of the brain: a personal mode and a meta-reflective mode. The personal mode refers to a motivationally and socially embedded configuration oriented toward adaptation, identity maintenance, and ego-relevant concerns. The meta-reflective mode emerges when cognition turns upon itself and becomes its own object, enabling abstraction, self-objectification, and existential evaluation. The model does not posit metaphysical dualism nor anatomically isolated systems; rather, it interprets both modes as dynamically shifting configurations within overlapping neural

 doi:10.20944/preprints202603.0446.v1

substrates. The aim of this paper is not to provide empirical verification but to articulate a coherent structural framework capable of integrating phenomenological observations with contemporary neurocognitive perspectives. By analyzing the interaction and tension between these functional states, the model seeks to clarify distinctions between different categories of psychological crises and to contribute to ongoing discussions in neurophilosophy regarding the organization of conscious experience. Broader cultural or social implications remain exploratory and are presented as potential extensions rather than definitive conclusions.

Main

1. The Formation and Structure of the Personal Mode

The personality mode is an evolutionarily established configuration of the human brain through which biological survival mechanisms gradually transform into structured self-representation. Its origins lie in instinctive regulatory systems governing threat detection, reward anticipation, attachment, and dominance. Over approximately the last million years of hominin evolution, particularly with the gradual expansion of the prefrontal cortex and the strengthening of long-range cortical connectivity, these affective systems became integrated with memory, symbolic processing, and future-oriented planning. The result was not the suppression of instinct, but its reorganization into a coherent and narratively stabilized self-model.

Early hominins relied primarily on subcortical and limbic circuits that rapidly assessed environmental stimuli in terms of survival relevance. However, as associative cortical regions expanded and frontal areas developed greater regulatory capacity, emotional valence and motivational significance became embedded within increasingly complex cognitive structures. Autobiographical memory, long-term social positioning, and anticipatory simulation gradually integrated into a unified perspective organized around personal continuity. The outcome was not merely a reactive organism, but a self-referential system capable of representing itself across time.

Within this architecture, midline cortical structures, and particularly the ventromedial prefrontal cortex (vmPFC), play a central integrative role. The vmPFC is deeply involved in valuation, emotional meaning attribution, autobiographical integration, and identity coherence. Through dense interconnections with limbic regions, the hippocampus, and broader associative networks, it contributes to the stabilization of a relatively stable self-model. This model links past experience, present evaluation, and anticipated future into a unified structure of relevance. The “self” within this framework is not a metaphysical entity, but a dynamically maintained integrative model weighted by value.

The personality mode is further modulated by endocrine dynamics. Hormonal states alter the architecture of perception itself. Elevated cortisol narrows interpretative framing toward heightened threat sensitivity and risk anticipation. Fluctuations in dopaminergic activity influence motivational salience and future projection. Oxytocin modulates group affiliation and trust attribution. Testosterone shifts competitive and dominance-related evaluation. These biochemical regulators influence more than subjective feeling; they alter which aspects of reality become foregrounded and which recede. Thus, the perceived world is partially shaped by neurochemical configuration at any given moment. At the same time, the personality mode retains regulatory control, while hormonal activity functions largely as an amplifier of already established value structures.

Modern cognitive science increasingly conceptualizes the brain as a predictive modeling system. Perception is not passive reception of external data, but active construction guided by prior expectations and value hierarchies. Sensory input is interpreted through internally generated models that prioritize coherence with existing self-representations. In this respect, the personality mode does not engage reality neutrally; it organizes reality around the preservation and continuity of the self-model. This structural mediation resonates with philosophical insights articulated by Aristotle, who argued that cognition operates through organizing principles inherent to the perceiving subject rather than through apprehension of objects in absolute purity. Contemporary neuroscience

 doi:10.20944/preprints202603.0446.v1

reformulates this insight mechanistically: the brain constructs adaptive representations rather than directly reflecting objective reality. Although humans have developed increasing awareness of their own motivational structures, the genetic and biological foundations of personality remain inextricably involved.

Throughout evolution, especially in Homo sapiens, this self-integrative architecture became progressively more abstract. Systems initially tuned to immediate tribal survival expanded to encode symbolic roles, long-term aspirations, moral commitments, and complex social hierarchies extending far beyond direct perception. The modern personality mode is therefore not a primitive survival residue, but a culturally reinforced and symbolically expanded identity structure. It integrates biological drives with narrative coherence, future modeling, and social positioning. The self becomes a long-term project rather than a momentary adaptive state. Due to the relative expansion of the prefrontal cortex compared to other primates, personal goals became increasingly abstract. This abstraction enabled prioritization of internal development and self-directed growth, which in turn facilitated the emergence of metacognitive capacities during cultural evolution. Nevertheless, personality and its biologically grounded motivational core remained fundamental. Despite the unprecedented expansion of metacognition and self-reflection in humans, personality retained structural priority, influencing thought processes, information processing, and perceived objectivity.

However, the same integration that enables personal stability introduces epistemic constraints. Because perception and evaluation are filtered through value-laden predictive models centered on self-coherence, information inconsistent with identity commitments may be reinterpreted or attenuated. The brain preferentially maintains internal narrative stability over representational neutrality. Consequently, what appears as purely rational evaluation often emerges from prior affective weighting embedded in the personality mode.

Meta-reflective cognition—the reasoned examination of one's own beliefs and assumptions—does not operate outside this architecture. Although dorsolateral prefrontal systems contribute to abstract reasoning and cognitive control, they remain interconnected with valuation networks centered in regions such as the vmPFC. Reflection therefore evaluates representations that have already been shaped by self-referential filtering. The mind that seeks objectivity does so using instruments shaped for adaptive coherence rather than detached neutrality. The personality mode thus constitutes the foundational configuration of human cognition: an evolutionarily shaped, hormonally modulated, neurally integrated system that constructs a self-centered world-model. It enables identity, continuity, and social complexity, yet simultaneously limits the possibility of fully unconditioned meta-cognitive perception. The tension between personal coherence and rational examination is not incidental, but structurally embedded within the architecture of the human brain.

2. The Formation of the Rational Mode: Biological and Cultural Emergence of Metacognition

The rational mode, understood here as meta-reflective cognition, represents not a separate ontological faculty but a functional reconfiguration of pre-existing neural architecture. If the personality mode stabilizes a coherent self-model, the rational mode introduces the capacity to represent that model as an object of analysis. This shift—from self-representation to representation of representation—marks a critical threshold in cognitive evolution. The biological foundations of this transition are closely linked to the relative expansion and increased functional specialization of the prefrontal cortex in Homo sapiens. While the ventromedial prefrontal cortex (vmPFC) contributes to valuation and identity coherence, the dorsolateral prefrontal cortex (dlPFC) plays a central role in working memory, abstract reasoning, cognitive control, and hierarchical rule integration. The dlPFC enables the temporary suspension of immediate affective impulses and supports manipulation of symbolic structures independent of direct survival relevance. This capacity allows cognition to operate on its own contents. Importantly, this development did not create a new "substance" of mind. Rather, it reorganized existing predictive and self-referential systems into more complex recursive loops. Through strengthened fronto-parietal networks and enhanced long-range connectivity, the

 doi:10.20944/preprints202603.0446.v1

brain became capable not only of understanding internal states but of recognizing that it understands them. Metacognition emerges when the brain models its own modeling activity.

At the same time, the personality mode—although not merely a defensive reflex—remains structurally foundational. Through its integration with self-preservation mechanisms and socially embedded value frameworks, it can buffer individuals against existential disorientation. By embedding meaning in social roles, attachments, goals, and even minor achievements, the personality mode protects against experiences of radical meaninglessness. It allows individuals to derive satisfaction from everyday structures and localized purposes. In this sense, it stabilizes psychological continuity not through denial of existence, but through selective anchoring of relevance.

However, this same architecture can generate disproportionate suffering in domains such as intimate relationships or social positioning. Even when meta-reflective cognition identifies the triviality or irrationality of a given conflict—or even perceives a clear solution—the personality mode may continue to generate distress. The persistence of such reactions illustrates that cognition remains hierarchically grounded in the value-weighted self-model. Metacognition may reinterpret, but it does not abolish the underlying evaluative structure. This asymmetry further supports the claim that the rational mode operates as a superimposed regulatory layer rather than as a dominant foundation.

The full elaboration of meta-reflective thought required not only biological expansion but cultural accumulation. Language, writing, institutional memory, and intergenerational transmission created external scaffolding for recursive cognition. The emergence of philosophy in ancient civilizations constituted what may be described as the first wave of systematic self-examination(pre-modernism). Reflection became formalized as a disciplined inquiry into knowledge, existence, and reasoning itself. Later developments in physics and biology introduced a second wave: humanity began to understand itself through third-person explanatory frameworks (modernism). Scientific modeling allowed humans to construct accounts of their own biological constitution and cosmic position, thereby approximating a view of the self from an externalized perspective. Yet even this third-person vantage point remains internally generated and neurally mediated.

Neurophilosophical approaches emphasize precisely this continuity between biological structure and reflective reasoning. Thinkers such as Patricia Churchland argue that philosophical analysis must remain continuous with neuroscience, rejecting strict dualisms between reason and biology. From this standpoint, metacognition does not transcend neural constraint; it refines and reorganizes it. Crucially, the rational mode is metabolically and computationally demanding. Sustained activation of dlPFC-mediated control networks requires significantly greater energetic expenditure than default affective processing. Abstract reasoning, inhibitory control, perspective-shifting, and recursive modeling impose measurable neural costs. Consequently, meta-reflective cognition functions as an energetically expensive superstructure rather than a constant baseline state. Under stress, fatigue, or strong emotional activation, the system naturally reverts toward the more stable and evolutionarily older personality configuration.

The rational mode should therefore be understood as a biologically grounded, culturally amplified, and energetically costly extension of the personality mode rather than an independent entity. It depends on the same valuation systems and predictive architectures that structure personal identity. Even when attempting to detach from bias, reflective reasoning recruits networks embedded within the broader self-integrative system. The capacity to "see oneself from the outside" remains biologically situated.

Thus, while cultural evolution expanded recursive depth—enabling philosophy, science, and existential inquiry—the underlying hierarchy persists. The personality mode provides motivational grounding, affective weighting, and meaning stabilization. The rational mode introduces recursive examination and structural abstraction. The tension between them is not accidental but inherent: reflection may reinterpret the self, yet it remains anchored in the very architecture it seeks to evaluate.

3. Interaction, Conflict, and Integration of Personality and Meta-Reflective Modes

 doi:10.20944/preprints202603.0446.v1

The personality and meta-reflective modes, while functionally interdependent, often exhibit divergent priorities and evaluative frameworks, producing bot tension and complementarity in human cognition. The personality mode organizes perception, memory, and motivation around self-coherence, social embeddedness, and adaptive salience. In contrast, the rational mode introduces recursive evaluation, abstract reasoning, and meta-reflective scrutiny. This distinction generates differences in how the same stimuli or situations are interpreted. For example, a social conflict may elicit immediate distress and concern for status or relational stability in the personality mode, while the rational mode may recognize the situation as trivial or strategically manageable. The divergence illustrates that human experience is rarely unitary: multiple interpretive layers operate simultaneously, sometimes in alignment, sometimes in tension.

These differences are decreasing when the issue of mortality is introduced. The prospect of death disrupts the habitual organization of the personality mode, producing existential uncertainty that exceeds routine evaluative frameworks. The meta-reflective mode, by modeling time, causality, and abstract continuity, can contextualize mortality in broader conceptual systems, offering explanatory or symbolic structures. However, even when the rational mode(meta-reflective) identifies strategies for existential reconciliation, the affective and evaluative tendencies of the personality mode often persist. The result is a partial dissolution of the boundary between modes: the experience of mortality can simultaneously engage deeply personal affective response and high-order reflection, blurring the distinction between self-centered concern and detached analysis.

Manifestations of mode dominance vary across individuals and contexts. When the personality mode is strongly active, cognitive and affective processing is guided primarily by self-relevant evaluation, social salience, and motivational coherence. In these cases, meta-reflective reasoning may be attenuated, yielding behavior that is adaptive in social or survival contexts but less capable of abstract problem-solving. Conversely, in individuals with highly engaged rational modes, abstract reflection and strategic analysis can dominate attention and decision-making. Even in such cases, however, the personality mode remains partially operative: residual affective priorities, habitual valuation, and socially embedded motivations continue to shape perception and judgment, preventing complete detachment from the self. This partial persistence explains why even those closely aligned with meta-reflective cognition retain affective and social vulnerabilities and why ego-centered considerations never fully disappear. Suppression and amplification of modes can occur dynamically. Stress, fatigue, or immediate threat often transiently elevate personality-driven processing, whereas structured reflection, education, or meditation may enhance rational mode engagement. Excessive dominance of either mode carries characteristic limitations. Overactive personality mode can produce rigidity, overidentification with self-relevant narratives, and emotional bias. Overactive rational mode, conversely, may generate detachment, analytical paralysis, or neglect of social and affective cues. Optimal cognitive functioning arises from a flexible interplay: both modes remain active but contextually weighted according to demands and capacities. Neurophilosophical perspectives reinforce this dynamic. While no single cortical or subcortical area uniquely “houses” a mode, networks involving prefrontal, cingulate, and associative regions contribute to hierarchical evaluation and affective regulation. Recursive loops between valuation, memory, and cognitive control networks underpin the capacity to switch between modes and maintain partial integration. From a Patricia Churchland–inspired perspective, the modes are not metaphysically separate: they are emergent functional configurations of a continuous neural system. Their interaction, tension, and integration constitute a phenomenologically real architecture, albeit one embedded in biology and shaped by social and cultural experience.

The interplay of modes becomes particularly salient when decisions require reconciliation of long-term goals with immediate affective imperatives or existential concerns. Personality-driven evaluation often anchors cognition in socially and biologically salient priorities, while rational reflection permits abstract assessment of meaning, mortality, and strategy. In practice, human experience is a dynamic negotiation between these layers. Even when meta-reflective reasoning is highly active, it is constrained by the foundational architecture of personality. Conversely,

 doi:10.20944/preprints202603.0446.v1

personality mode activity can be moderated, reframed, or temporarily subordinated by rational oversight, yet never entirely eliminated. This persistent asymmetry underscores the hierarchical integration central to human cognition: the rational mode extends and refines, but ultimately remains anchored in, the personality mode. In sum, personality and rational modes are complementary yet distinct configurations whose interaction underlies human adaptability, reflection, and existential engagement. Their divergence in evaluative priorities, capacity for integration, and partial dominance provides both the flexibility and constraint inherent to human experience. Confrontations with mortality, stress, or complex social dynamics highlight the dynamic balance between modes, revealing the

4. Social Influence and the Formation of Collective Human Character

The dynamics of personality and rational modes, when extended across populations, provide a conceptual framework for understanding the emergence and structure of civilization. Individual cognition does not operate in isolation: humans are biologically and culturally embedded within a species-wide network of interaction, communication, and shared practices. From a functional perspective, this gives rise to what may be described as a collective biological personality—an emergent configuration reflecting the general tendencies, motivations, and affective patterns common across Homo sapiens.

The collective personality is neither metaphysical nor abstract in the Hegelian sense of the "Absolute Spirit"; it is not an idealized universal consciousness. Rather, it is grounded in shared biological and social predispositions: a broadly distributed set of affective priorities, curiosity-driven exploration, and what may be termed an almost instinctive drive toward social engagement. These commonalities form the substrate for stable social structures, norms, and coordinated action. Human tendencies toward cooperation, information sharing, and cumulative cultural learning exemplify the operational expression of the collective personality. It is through this shared architecture that psychology and sociology are able to identify generalizable patterns: without it, interindividual variability would render systematic study intractable, and cross-cultural generalizations would lack coherent foundation.

The personality mode of the individual interacts with this collective substrate by selectively amplifying, attenuating, or reorganizing tendencies inherited from shared human predispositions. While individual personality embodies a unique configuration of affective weighting, social experience, and motivational priority, the underlying collective architecture ensures a baseline coherence of behavior and cognition. In this sense, civilization can be understood as a large-scale stabilization of collective tendencies, structured by both biological constraints and cultural scaffolding. The rational mode introduces a further layer of differentiation. Metacognition enables individuals to step outside the immediate orientation of the collective personality, reflecting on social norms, questioning inherited practices, and experimenting with alternative value structures. In doing so, rational cognition imposes constraints on the collective personality, producing variation and innovation. Yet even in the presence of highly active metacognitive processes, residual traces of collective personality persist: foundational drives for curiosity, communication, and social embeddedness remain influential, shaping perception, motivation, and behavior. This interplay reflects the hierarchical architecture of cognition articulated in prior sections, and echoes the Aristotelian observation that all experience is mediated through the organizing principles of the perceiver, rather than unfiltered reality.

Historical and contemporary patterns of civilization illustrate this dual influence. Large-scale cooperation, technological innovation, and cultural accumulation are possible only because the collective personality provides shared motivational and cognitive frameworks. Simultaneously, individual rationality enables the critical evaluation of inherited norms, the creation of science, philosophy, and institutions, and the adaptive modification of collective practices. The persistence of the collective personality within highly metacognitive individuals underscores the embeddedness of

 doi:10.20944/preprints202603.0446.v1

all human cognition within biologically and socially constrained networks, even as rational reflection enables conceptual distance and novelty.

In summary, social influence and collective human character emerge from the interaction of shared biological predispositions and individual cognitive architectures. Civilization is built upon the stabilization of collective tendencies, while rational reflection introduces divergence, innovation, and critique. The tension between the collective and the individually reflective modes mirrors the structural hierarchy of cognition itself: the personality mode forms the substrate, the rational mode extends and constrains, yet neither can fully abolish the other. This perspective situates human social life within a neurophilosophically grounded framework, where both biological commonality and meta-reflective individuality are necessary for understanding the formation, persistence, and evolution of complex societies.

5. Religion as the Product of Metacognitive Demand

Religion can be understood not merely as a cultural artifact or as a manifestation of fear of death, but as a functional expression of metacognitive necessity. While early human experiences of mortality initially generated anxiety within the personality mode, the rational mode—capable of recursive reflection on existence—produced the cognitive demand to stabilize meaning beyond immediate survival concerns. Religion, in this framework, functions as a neurocognitively grounded system: it provides a structured narrative and symbolic scaffolding that mitigates existential uncertainty, offering a persistent "background meaning" to dampen the noise of mortality awareness.

The development of religious systems illustrates how metacognition interacts with both individual personality and collective human tendencies. The collective biological personality, while essential for social cohesion, exhibits minimal intrinsic demand for meta-reflective frameworks. Its focus lies in curiosity, communication, and cooperative survival; these tendencies are sufficient for establishing basic social structures, but they do not inherently generate abstract reflection on ultimate meaning. Consequently, the early role of religion was to satisfy a metacognitive need and the desire of the personal mod to quickly return to the dominance of the collective personality to provide interpretive order to the contingencies of existence, thereby enabling individuals to engage in social coordination without being paralyzed by existential uncertainty.

At the individual level, the personality mode can facilitate the rapid stabilization of collective structures. Heidegger's notion of Dasein exemplifies this dynamic: the embodied, socially situated self tends to prioritize engagement with practical, lived concerns. By anchoring attention in immediate social, affective, and goal-directed tasks, the personality mode can suppress excessive meta-reflective rumination. This suppression, paradoxically, allows the collective personality to organize effectively, constructing institutions, norms, and cultural frameworks—including religious systems—without being overburdened by abstract existential anxiety. In this way, religion emerges at the intersection of metacognitive necessity and personality-driven structuring: it channels reflective inquiry into socially and cognitively tractable forms.

Over historical time, this dynamic expanded to broader sociocultural structures. As civilizations grew, religion initially provided coherent narratives of order, morality, and teleology. Later, other large-scale collective projects, such as nation-states and political ideologies, assumed similar functional roles, distributing meta-reflective demands across institutions rather than relying solely on individual cognition. In each case, the personality mode continues to mediate: by integrating with socially normative pressures and the influence of Das Man, it suppresses excessive metacognitive rumination, allowing collective personality structures to stabilize efficiently. Religion and analogous institutional frameworks thereby enable the collective biological personality to organize and persist while maintaining coherence across large populations.

The personality mode, through its interplay with Das Man, can effectively anchor cognition in immediate social and practical concerns, providing continuity and stability. By moderating the expression of the rational mode, it prevents the destabilizing effects of constant existential reflection, allowing collective structures—rituals, norms, and institutions—to form robustly. Metacognition,

 doi:10.20944/preprints202603.0446.v1

while constrained, is redirected toward symbolic elaboration, normative critique, and refinement rather than paralyzing individual or social action.

In summary, religion represents a neurophilosophically interpretable product of the interaction between meta-reflective cognition and personality-mode dynamics. It arises not solely from fear of death, but from the rational mode's demand to stabilize meaning in the face of existential uncertainty. The collective biological personality lacks intrinsic meta-reflective requirements, so religion functions to scaffold both individual understanding and social cohesion. Through the moderating influence of the personality mode within the context of Das Man, collective structures can emerge efficiently, demonstrating the hierarchical and interactive architecture of cognition in the shaping of civilization

6. Crises of Personality and Cognition: Distinguishing Biological and Metacognitive Conflicts

Human experience is punctuated by crises that reflect the interplay between personality and rational modes, yet these crises differ fundamentally in their origin, scope, and phenomenology. Age-related, intra-personal crises—such as the classic "midlife crisis"—emerge primarily from the personality mode. These crises are rooted in affective evaluation, social positioning, biological rhythms, and the regulatory mechanisms that stabilize the self-model over time. They reflect challenges to identity continuity, social status, and motivational coherence, rather than failures of abstract reflection or meta-cognitive processing.

Midlife crises, adolescence-related conflicts, and similar developmental perturbations are largely biologically and socially mediated. Hormonal fluctuations, neural plasticity, and the accumulation of social feedback generate periods of tension where the personality mode reassesses life goals, relational networks, and personal achievements. These episodes are experienced as highly salient, emotionally charged, and contextually constrained, yet they do not necessarily implicate meta-reflective reasoning. In other words, the distress arises from the valuation and affective weighting intrinsic to personality, not from a failure to model or reflect upon one's own cognition. In contrast, existential crises—moments of profound awareness of mortality, finitude, and the apparent contingency of meaning—originate primarily within the rational mode. Such crises engage recursive reflection, abstract temporal modeling, and evaluation of the self from a third-person perspective. The anxiety they provoke is not simply a reaction to immediate social or biological circumstances; it arises from the rational mode's recognition of structural contingencies, uncertainty, and the limits of personality-generated meaning. The reflective evaluation of existence thus produces a qualitatively different phenomenology: it is more abstract, temporally expansive, and epistemically oriented.

These distinctions have important implications for understanding human development and well-being. Age-related crises are nearly universal, predictable, and structurally embedded in biological and social mechanisms. They can be observed across cultures, and they follow patterns consistent with maturation, social role acquisition, and hormonal changes. By contrast, existential crises are contingent upon the development and engagement of metacognitive faculties; they require reflective awareness of self and temporality. Not all individuals experience existential anxiety with the same intensity or frequency, highlighting the dependence of such crises on rational-mode activation.

The interaction of these modes during crisis periods also elucidates potential conflicts. Personality-mode crises can be partially mitigated by rational evaluation, yet the underlying affective salience persists. Similarly, existential reflection may be constrained by the personality mode's prioritization of adaptive, socially embedded concerns, resulting in partial suppression of radical metacognitive anxiety. This asymmetry reinforces the hierarchical organization of cognition: while meta-reflective processes can reinterpret or contextualize experience, they remain anchored in the biologically grounded evaluative framework of the personality mode.

In sum, the distinction between age-related personality crises and metacognitive existential crises underscores the functional and phenomenological separation of the two cognitive modes. Understanding the biological and social foundations of personality-mode crises illuminates why such experiences are predictable and broadly shared, while existential crises highlight the unique, energy-

 doi:10.20944/preprints202603.0446.v1

intensive demands of metacognitive engagement. Together, these patterns reveal the layered architecture of human cognition, in which evolutionarily grounded personality dynamics coexist with culturally and biologically amplified reflective capacities.

7. Cultural Units (Memes) and the Neurocognitive Conflict Between Biological Structures and Metareflection

The interaction between personality and rational modes does not unfold in a culturally neutral environment. Beyond genetic predispositions and individual development, cognition is continuously shaped by transmissible cultural units—memes—which stabilize and replicate patterns of interpretation, valuation, and response across individuals. In the sense introduced by Richard Dawkins, memes function as units of cultural transmission. Within the present framework, however, they may be interpreted more precisely as distributed regulators of cognitive mode dominance.

As outlined in Sections 1 and 3, the personality mode rests upon evolutionarily older, energetically efficient neural architectures that prioritize threat detection, coalition formation, status evaluation, and immediate affective coherence. The rational or metacognitive mode, in contrast, is developmentally later, metabolically more demanding, and dependent upon extended associative and regulatory networks. It enables recursive self-modeling, abstraction, and tolerance of ambiguity, yet it does so at increased energetic and temporal cost. This asymmetry establishes a structural tension: under conditions of uncertainty or stress, cognitive processing tends to reweight toward older, biologically stabilized configurations.

Cultural units can amplify or attenuate this reweighting. Certain meme structures—particularly those characterized by simplified moral binaries, unquestionable symbolic authority, and repetitive narrative reinforcement—reduce the necessity for recursive evaluation. They provide ready-made interpretive frameworks that satisfy the personality mode's need for coherence and social alignment, thereby decreasing the engagement of metacognitive processing. This modulation does not eliminate rational capacity but constrains its operational scope. Under high-threat conditions, such as warfare, these meme configurations can stabilize rapid coalition alignment and moral clarity, reinforcing older biological priorities while minimizing the energetic burden of sustained reflection.

Conversely, other meme complexes—such as institutionalized scientific method, philosophical skepticism, and norms of open critique—scaffold the rational mode. They legitimize doubt, formalize delayed judgment, and distribute the cognitive load of reflection across social structures. In this way, civilization may be understood as partially externalizing metacognitive effort into institutions, allowing reflective practices to persist despite their energetic cost. These cultural forms do not replace biological substrates but create environments in which metareflection is sustained rather than suppressed. The developmental trajectory of individuals further illustrates this interaction. Genetic predispositions bias temperament, sensitivity to threat, reward processing, and impulsivity. Early childhood memes—language structures, parental narratives, religious or national identity frameworks—are internalized during periods of high neural plasticity. These early-embedded patterns shape the baseline configuration of the personality mode before metacognitive capacities are fully matured. As a result, later-developed rational reflection operates upon a pre-structured evaluative system. Metacognition may reinterpret or critique these inherited configurations, yet it rarely abolishes them entirely. This developmental layering reinforces the asymmetry between foundational biological architectures and subsequent reflective extension.

The metaphor of "war" between older biological systems and metacognition, should be interpreted functionally rather than ontologically. There is no discrete anatomical boundary separating these configurations; rather, distributed neural networks are dynamically weighted according to context, energy availability, and cultural reinforcement. Under conditions of chronic stress or collective threat, biologically conserved processing strategies gain relative dominance. Under stable, resource-rich conditions supported by reflective institutions, metacognitive engagement becomes more sustainable. Memes thus participate in shifting the balance of this dynamic system.

 doi:10.20944/preprints202603.0446.v1

Importantly, the rational mode remains dependent upon the motivational substrate of the personality mode. Curiosity, social engagement, and affective salience—features of the collective biological personality described in Section 4—provide the initial impetus for reflective inquiry. Metacognition, although evolutionarily later and energetically costly, does not stand outside biology; it is an extension layered upon ancient neural structures. Its apparent distance from biological immediacy reflects temporal depth in evolutionary history and metabolic demand, not metaphysical separation.

8. Existentialism and Its Limits: Biological Constraint, Metacognitive Boundaries, and Artificial Extension

Existentialist thought, particularly in its twentieth-century formulations, emphasizes radical freedom, self-construction, and the primacy of subjective responsibility. The individual is often described as fundamentally self-defining, projecting meaning onto an otherwise indifferent world. While this perspective captures an important dimension of metacognitive capacity—namely, the ability to evaluate and reconstruct one's own interpretive frameworks—it risks overstating the degree of structural independence available to human cognition.

Within the present model, the individual is not a blank slate. Each person emerges as a biologically instantiated system structured by genetic predispositions, early developmental constraints, and culturally transmitted memes acquired during periods of heightened neural plasticity. Temperament, affective thresholds, reward sensitivity, and stress reactivity are not freely chosen; they are partially encoded in genetic architecture. Early language acquisition, familial narratives, religious or national identity structures, and social reinforcement patterns further shape the baseline configuration of the personality mode before the rational mode reaches full operational maturity.

Thus, existential freedom operates within a pre-structured cognitive field. Metacognition can reinterpret, reweight, and reorganize these inherited structures, but it does not arise in isolation from them. The collective biological personality described previously—characterized by curiosity, social embeddedness, and an almost instinctive orientation toward communication—provides a shared substrate across individuals. This collective dimension ensures that even the most radically self-reflective agent remains embedded within species-typical motivational and perceptual architectures. Existential self-creation is therefore constrained by biological and social continuity. The limits of existentialism become particularly visible when examining the boundaries of metacognitive capacity. Metareflection is energetically expensive, developmentally late, and structurally dependent on underlying affective and motivational systems. It enables recursive modeling of the self, abstraction from immediate valuation, and third-person evaluation of one's own cognitive processes. However, it cannot fully detach from the personality mode upon which it depends. Attempts at radical detachment often encounter affective resistance, motivational depletion, or cognitive fatigue. The system cannot sustain indefinite recursive abstraction without destabilizing the coherence maintained by personality structures.

In extreme forms, excessive metacognitive engagement may approach self-undermining tendencies. Continuous recursive doubt can erode motivational salience; hyper-reflection may weaken identity coherence; persistent abstraction can attenuate embodied engagement with lived experience. These outcomes do not represent metaphysical failure but functional limits: a biologically embedded system cannot indefinitely prioritize high-cost reflective operations without rebalancing toward affective and social grounding. The potential "self-destruction" of metacognition, therefore, lies not in annihilation but in destabilization of the hierarchical integration between modes. At the civilizational scale, this tension is mediated by institutional and cultural scaffolding. Philosophical systems, scientific methodology, and reflective traditions distribute the cognitive burden of metareflection across collective structures, preventing individual overload. Nevertheless, even the most advanced reflective cultures remain anchored in the collective biological personality. Norms, identities, and emotionally salient narratives continue to provide motivational structure beneath formal rationality.

 doi:10.20944/preprints202603.0446.v1

The emergence of artificial intelligence introduces a novel dimension to this architecture. AI systems can extend certain functions of the rational mode: large-scale pattern detection, abstract modeling, probabilistic inference, and even forms of simulated meta-representation. In doing so, they may augment human metacognitive capacity by externalizing analysis and expanding informational scope. However, such systems remain constrained by the informational substrate from which they are trained. They are constructed from data generated by a single biological species and reflect its linguistic, conceptual, and cultural distributions.

Consequently, AI does not transcend the structural limits of human cognition; it amplifies and reorganizes them. It lacks the embodied affective grounding characteristic of the personality mode and does not participate in the collective biological personality except indirectly, through the data it processes. While it can expand the operational reach of metareflection—allowing more complex simulations and broader informational integration—it remains epistemically bounded by the representational patterns of Homo sapiens. Its apparent objectivity is therefore derivative rather than ontologically independent.

In conclusion, existentialism captures a genuine dimension of human metacognitive capacity: the ability to reinterpret and restructure meaning. Yet this freedom is constrained by genetic predispositions, early meme embedding, and the collective biological personality that shapes species-wide cognitive architecture. Metacognition itself has functional limits, dependent on energetic resources and affective grounding, and may destabilize if excessively detached from its biological substrate. Even technological extensions such as artificial intelligence expand rather than abolish these constraints, operating within the informational horizon of the human species. Human cognition thus remains a layered system: biologically grounded, culturally modulated, and reflectively extensible, yet never fully unbounded.

Discussion

1. What Has Been Established

The present work has not merely proposed a distinction between personality and metacognitive configurations; it has clarified their hierarchical interdependence and structural asymmetry. What emerges is not a dualistic separation, but a layered cognitive architecture in which biologically grounded evaluative systems provide the substrate upon which reflective extension becomes possible.

First, it has become evident that personality is not a superficial psychological construct but a biologically stabilized integrative regime. It organizes affect, motivation, social orientation, and identity coherence in ways that are evolutionarily conserved and metabolically efficient. Personality is not a defensive illusion nor an accidental narrative overlay; it is the primary organizing structure of human cognition. Even when reflective capacities become highly developed, they do not replace this substrate. Rather, they operate upon it.

Second, metacognition has been clarified as an energetically costly, developmentally late, and culturally scaffolded extension of this architecture. It enables recursive modeling—understanding not only external reality but the structure of one's own understanding. This recursive capacity permits abstraction, third-person perspective taking, and the construction of philosophy and science. However, it does not exist independently of the personality mode. Its operations are constrained by affective weighting, motivational salience, and early-embedded cognitive patterns.

Third, the model demonstrates that civilization itself reflects the interaction between these regimes. Collective biological personality—shared tendencies toward curiosity, communication, and social embedding—provides the baseline coherence that makes large-scale structures possible. Religion, national identity, and ideological systems can be interpreted as cultural stabilizers that manage the energetic tension introduced by metacognitive awareness of mortality and contingency. At the same time, scientific and philosophical traditions represent institutional scaffolding for sustained reflective engagement.

 doi:10.20944/preprints202603.0446.v1

Fourth, developmental asymmetry has been clarified. Genetic predispositions and early memes shape the evaluative core of personality before metacognition reaches maturity. This layering explains why reflective reasoning rarely abolishes deeply internalized commitments. Existential freedom, therefore, operates within constraints: humans are not blank slates but structured systems capable of reinterpretation rather than total reconstruction.

Fifth, the limits of metacognition have become conceptually visible. Because reflective processing is metabolically expensive and dependent on underlying motivational systems, it cannot sustain indefinite recursive abstraction without destabilizing the coherence maintained by personality. Excessive detachment risks fragmentation of meaning and motivational depletion. Thus, the aspiration toward radical existential autonomy encounters biological boundaries.

Finally, the introduction of artificial intelligence highlights an externalization of rational-mode functions. AI may extend analytical and integrative capacities, yet it remains epistemically bounded by human-generated data and lacks participation in the collective biological personality. It expands metareflection but does not transcend the structural constraints of the species from which its informational substrate derives.

In sum, what has been established is a coherent interpretation of human cognition as a hierarchical, biologically anchored system in which personality constitutes the foundational integrative regime and metacognition operates as an extension rather than an independent essence. Civilization, existential reflection, religion, crisis phenomena, and cultural evolution can all be reinterpreted within this layered framework. The resulting model does not eliminate philosophical freedom, but it situates it within structural limits imposed by biology, development, and collective continuity.

2. Limitations of the Article

Despite the internal coherence of the proposed framework, several limitations must be acknowledged. First, the article remains primarily conceptual. It synthesizes findings from cognitive science, evolutionary theory, and philosophy of mind, but it does not introduce new empirical data. The distinction between personality as an integrative biological regime and metacognition as a reflective extension is theoretically motivated rather than experimentally isolated within a controlled framework. Contemporary research on metacognition—such as work by Stephen M. Fleming and Hakwan Lau—demonstrates measurable neural correlates of confidence monitoring and introspective accuracy, yet these findings do not straightforwardly validate a two-regime structural interpretation. The present model extrapolates from such findings rather than directly testing its architectural claims.

Second, the proposed division between personality and metacognitive modes risks oversimplification. Neural systems underlying valuation, self-representation, and executive control are deeply interconnected. Research associated with Michael Gazzaniga and large-scale brain network models suggests distributed integration rather than clean functional partitions. Therefore, the distinction articulated in this article should be interpreted as a heuristic abstraction rather than a strict neuroanatomical demarcation.

Third, the treatment of genetic predispositions and early memes as structural constraints, while consistent with evolutionary developmental perspectives, lacks operational specification. The concept of the meme, introduced by Richard Dawkins, remains theoretically influential but empirically diffuse. Cultural transmission is measurable, yet its direct mapping onto stable personality configurations requires more rigorous methodological development than is provided here.

Fourth, the notion of a collective biological personality is interpretive. While evolutionary psychology and social neuroscience support the existence of species-typical social drives, describing them as a coherent "collective personality" extends beyond established terminology. This construct serves explanatory integration but requires clearer definitional boundaries and measurable indicators to function as a scientific category.

 doi:10.20944/preprints202603.0446.v1

Fifth, the discussion of existential limits and self-destabilization through excessive metacognition remains partially philosophical. While clinical psychology documents phenomena such as depersonalization and derealization under conditions of cognitive overload or existential anxiety, the article does not provide systematic clinical data linking these states to the structural model proposed here. The connection is theoretically plausible but empirically underdeveloped.

Sixth, the analysis of artificial intelligence as an extension of reflective capacity is constrained by current technological conditions. Systems such as OpenAI language models operate through statistical pattern integration rather than embodied biological regulation. However, the long-term trajectory of AI architectures remains uncertain. The argument that AI is limited by the informational domain of its training data may require revision as multimodal and adaptive systems evolve.

Finally, the interdisciplinary breadth of the article introduces epistemic tension. Integrating philosophy, neuroscience, evolutionary theory, and cultural analysis risks conceptual stretching. While this breadth is necessary for the proposed synthesis, it also increases the possibility of category conflation or theoretical overextension.

In summary, the limitations of this article arise from its conceptual nature, abstraction level, and interdisciplinary scope. The model should therefore be interpreted not as a finalized explanatory system but as a structured hypothesis space—one that invites operational clarification, empirical testing, and philosophical refinement.

3. Alternative Literature and Contrasting Positions

In addition to the model developed in this article, a broad range of scientific and philosophical perspectives address related aspects of personality, metacognition, consciousness, and selfhood. These alternative approaches often differ in foundational assumptions, methodologies, and interpretations, and provide points of comparison that help situate the present model within the broader literature.

One influential position in cognitive neuroscience and philosophy of mind is heterophenomenology, as proposed by Daniel Dennett, which treats subjective reports as data to be explained rather than as authoritative accounts of internal states. This approach rejects the assumption that introspective self-reports necessarily reflect veridical inner experience and instead emphasizes third-person interpretation of behavior and neural correlates. Heterophenomenology challenges models that grant privileged epistemic status to introspective metacognition, arguing that self-awareness must be understood as an interpretive output of distributed processes rather than as a distinct internal mode of cognition.

Some contemporary research in metacognition highlights complexity and context dependence rather than structural duality. For example, reviews in Nature Reviews Psychology underscore that metacognitive mechanisms underlying confidence, self-reflection, and beliefs about one's own cognition are variable, influenced by task demands, and frequently altered in clinical conditions. These findings suggest that metacognition is not a unitary "mode" but rather a cluster of interacting processes that may not align cleanly with a hierarchical structural dichotomy.

Alternative neuroscientific perspectives emphasize embodied cognition, which situates cognitive capacities within sensorimotor interaction with the environment rather than within internal reflective architectures alone. Such approaches argue that cognition—including self-awareness—is shaped by ongoing organism–environment coupling, calling into question models that treat internal reflective mechanisms as primary. Embodied cognition frameworks maintain that body-world interaction plays a constitutive role in shaping cognitive patterns and that reflective abstraction emerges from, rather than overrides, these interactions.

In philosophy of consciousness, a wide variety of theories exist regarding the nature of self-reference and reflective awareness. The multiple drafts model, also associated with Dennett, denies a central "Cartesian theatre" of consciousness and instead views conscious experience as the result of parallel interpretive processes without a singular self locus. This perspective contrasts with models

 doi:10.20944/preprints202603.0446.v1

that posit structured self-representations as core cognitive constructs and raises questions about the ontological status of hierarchical cognitive modes.

Phenomenological and analytic traditions also present divergent views on self and metacognition. Some philosophers uphold that selfhood is fundamentally a narrative or relational construct rather than a biologically instantiated integrative system. For instance, arguments in the philosophy of consciousness emphasize minimal self or narrative self distinctions which situate selfhood in lived experience or social context, rather than as a stable neural architecture.

Integrated Information Theory (IIT) offers yet another approach by proposing that consciousness correlates with integrated information, conceptualizing conscious experience in terms of intrinsic causal power rather than hierarchical cognitive regimes. While IIT has been critiqued for suggesting panpsychist implications and for problematic interpretations of its own measures, it nonetheless diverges from mode-based architectures by prioritizing integrative capacity over functional separation of reflective systems.

Finally, methodological critiques from cognitive and personality psychology challenge the very comparability of personality theories. Meta-scientific analyses highlight that many traditional personality models differ significantly in definitions, operational criteria, and empirical foundations, indicating that any broad integrative theory of personality must engage with a heterogeneous empirical landscape.

In summary, alternative literature spans a spectrum from third-person cognitive frameworks, embodied and enactive paradigms, narrative and phenomenological accounts of selfhood, to formal theories of consciousness like IIT. These positions provide both complementary insights and critical challenges to hierarchical mode distinctions. Whereas the present model emphasizes structural interdependence of personality and reflective configurations, other frameworks either distribute self-referential functions across networks without hierarchical modes, reject introspective authority, or situate cognition in bodily and environmental interaction. Recognizing these alternative accounts encourages a broader, interdisciplinary dialogue and underscores the need for future work to integrate diverse approaches while acknowledging empirical constraints.

Conclusions

This article has proposed a conceptual framework for understanding human cognition as a hierarchical system in which personality constitutes the foundational, biologically anchored regime, and metacognition functions as a reflective, energetically costly extension. The model synthesizes insights from evolutionary theory, neuroscience, and philosophy to explain how affective, motivational, and self-referential structures interact with recursive reflective capacities.

Key findings of the model include:

The centrality of personality: Far from being a mere behavioral residue or adaptive reflex, personality organizes experience, emotion, motivation, and social positioning into a coherent self-model. It acts as the substrate for reflective processes and imposes structural constraints on metacognition. Personality also provides the mechanism through which humans navigate existential challenges, allowing meaning and satisfaction to be derived from everyday life, even under conditions of awareness of mortality.

The role of metacognition: Metacognitive processes—enabled primarily by dorsolateral prefrontal mechanisms and culturally scaffolded through philosophy, science, and symbolic systems—allow humans to observe, evaluate, and reinterpret their own mental states. Metacognition enables abstraction, third-person perspective taking, and the construction of systematic knowledge, but it is always contingent upon and constrained by the underlying personality structure. Excessive metareflexivity risks fragmentation of meaning and motivational depletion, highlighting the biological and cultural limits of reflective capacities.

Cultural and social implications: The interaction between personality and metacognition underlies the formation of collective human structures. Collective biological personality—shared drives for communication, curiosity, and social cohesion—establishes the baseline for civilization.

 doi:10.20944/preprints202603.0446.v1

Religion, national identity, and institutionalized knowledge systems can be understood as cultural stabilizers that negotiate the tension between personality-driven coherence and metacognitive awareness of contingency. Nevertheless, metacognition introduces individual uniqueness and limits uniformity, preserving a balance between collective continuity and personal reflection.

Constraints and future directions: Humans are neither blank slates nor purely rational agents; cognition is structured by evolutionary, developmental, and cultural constraints. Understanding these layered dynamics provides insight into crises, existential reflection, and the interplay of biology, culture, and reasoning. The emergence of artificial intelligence offers potential extensions of reflective capacities but remains bounded by human-generated informational substrates, illustrating the limits of meta-reflexive expansion outside the species-specific context.

In conclusion, the proposed conceptual model situates human selfhood, civilization, and reflective cognition within a structurally layered architecture. It emphasizes the integrative role of personality, the conditional extension of metacognition, and the interplay between biological, social, and cultural forces. This framework provides a platform for further theoretical development, empirical exploration, and philosophical inquiry, bridging the domains of neurocognitive science and philosophical reflection while acknowledging the limits inherent to both human cognition and methodological reach.

References

Churchland, P. S. (1986). Neurophilosophy: Toward a unified science of the mind-brain. MIT Press.

Churchland, P. S. (2013). Touching a nerve: The self as brain. W. W. Norton & Company.

Chalmers, D. J. (1996). The conscious mind: In search of a fundamental theory. Oxford University Press.

Dennett, D. C. (1991). Consciousness explained. Little, Brown and Company.

Dennett, D. C. (2017). From bacteria to Bach and back: The evolution of minds. W. W. Norton & Company.

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. Philosophical Transactions of the Royal Society B: Biological Sciences, 367(1594), 1338–1349. https://doi.org/10.1098/rstb.2011.0417

Fleming, S. M., & Frith, C. D. (Eds.). (2014). The cognitive neuroscience of metacognition. Springer. https://doi.org/10.1007/978-3-642-45190-4

Gazzaniga, M. S. (2011). Who's in charge? Free will and the science of the brain. HarperCollins.

Humphrey, N. (1992). A history of the mind. Chatto & Windus.

Lau, H. C., & Rosenthal, D. (2011). Empirical support for higher-order theories of conscious awareness. Trends in Cognitive Sciences, 15(8), 365–373. https://doi.org/10.1016/j.tics.2011.05.009

Nagel, T. (1974). What is it like to be a bat? The Philosophical Review, 83(4), 435–450. https://doi.org/10.2307/2183914