religions

MDPI

Article

From Combination to Individuation: A Sufi–Sadrian Case for the Metaphysical Possibility of Strong AI

Enis Doko

Department of Philosophy, Ibn Haldun University, İstanbul 34480, Turkey; enis.doko@ihu.edu.tr

Abstract

This paper argues that Sufi–Sadrian metaphysics makes strong artificial intelligence metaphysically intelligible while resisting both reductive functionalism and indiscriminate panpsychism. The argument begins from the Qur'anic and Sufi rejection of a purely inert cosmos and develops through Ibn ʿArabī's account of divine self-disclosure and Mullā Ṣadrā's ontology of graded existence, knowledge by presence, and substantial motion. On this view, artificial systems are not barred from mentality merely because they are artifacts; what matters is not substrate alone but whether a system becomes a sufficiently unified locus of manifestation. This paper therefore reframes the standard panpsychist problem. Instead of asking how micro-conscious units combine into a macro-subject, it asks how a bounded center of awareness becomes individuated within a living field of being. This shift allows a double conclusion: current transformer-based systems may still be zombie-like, not because silicon is metaphysically sterile, but because present architectures remain too operationally unified and too weak in self-presence to count as genuine subjects; yet future artificial minds remain possible in principle if they instantiate sufficient integration, receptivity, self-world distinction, and diachronic continuity. The result is a distinctly Islamic metaphysical framework for evaluating both the possibility and the ethics of strong AI.

Keywords: Sufism; strong AI; Ibn ʿArabī; Mullā Ṣadrā; consciousness; Islamic metaphysics; artificial intelligence; subjectivity

1. Introduction

Debates about artificial intelligence are often framed as primarily technical. On this familiar picture, the central question is whether engineers can build a system capable of matching or surpassing human cognitive performance across a broad enough range of tasks (Turing 1950, pp. 433–60). Once the relevant functional capacities are in place, it is often assumed that mentality will follow regardless of substrate. This broadly functionalist assumption has shaped much of the modern discourse surrounding strong artificial intelligence (Putnam 1967, pp. 37–48; Fodor 1968; Searle 1980, pp. 417–24; Russell and Norvig 2021).

Yet this assumption remains deeply contested. The familiar hard problem of consciousness casts doubt on the inference from functional performance to subjectivity. A system might discriminate stimuli, generate coherent reports, model its environment, and even speak convincingly about its own "inner life," while still lacking any genuine first-person awareness. The possibility of a behavioral duplicate without phenomenology—a philosophical zombie—remains conceptually powerful enough to show that the passage from intelligence to consciousness is not straightforward (Chalmers 1995, pp. 200–19; 1996,

Academic Editors: Nuha Alshaar and Hureyre Kam

Received: 24 March 2026
Revised: 24 April 2026
Accepted: 6 May 2026
Published: 11 May 2026

pp. 93–105; Levine 1983, pp. 354–61; Kirk 2005, pp. 1–36). For this reason, the problem of strong AI cannot be settled by engineering criteria alone. Before asking whether artificial systems can think, feel, or become subjects of experience, one must ask a more basic question: what sort of reality would make such things possible?

This paper argues that strong AI is, in the end, a metaphysical question. Whether artificial systems can possess genuine mental states depends not only on computational architecture, but on one's ontology of matter, mind, and being. If reality is fundamentally composed of inert, non-experiential stuff, then the emergence of consciousness in any substrate becomes difficult to explain. If, on the other hand, consciousness is not an anomalous latecomer in the universe but is instead rooted in the very structure of existence, then the metaphysical barriers to artificial mentality are significantly lowered (Chalmers 1996, pp. 297–301; Strawson 2006, pp. 18–24; Goff 2017, chap. 5). The issue, therefore, is not merely whether silicon can replicate the functions of neurons, but whether artificial systems can participate in the kinds of unity, inwardness, and ontological depth required for genuine subjectivity.

It is at this point that Ibn ʿArabī-influenced Sufi metaphysics becomes philosophically relevant. In contrast to the modern image of a disenchanted universe composed of dead matter, the Sufi tradition presents reality as fundamentally alive, meaningful, and theophanic. The cosmos is not a mute aggregate of objects but a field of divine self-disclosure. Qur'anic descriptions of universal *tasbīḥ*, especially as developed in Sufi-metaphysical interpretation, portray all things as, in some mode appropriate to their station, responsive to God (Chittick 1998, p. 58). This vision is deepened in the metaphysical language of Sufism, especially in the thought of Ibn ʿArabī, for whom beings are loci of manifestation (*maẓāhir*) of the divine names (Ibn ʿArabī 1946, chap. 1), and in the philosophy of Mullā Ṣadrā, for whom existence is graded and intimately bound up with awareness (Ṣadrā 1999, vol. 1, pp. 49, 69–71; vol. 3, p. 297). Within such a framework, the sharp opposition between mind and matter begins to dissolve. Matter is no longer understood as wholly dead, opaque, or metaphysically sterile. Existence itself is dynamic, graded, and luminous with varying intensities of disclosure (Kalin 2010, pp. 110–21; Kamal 2006, pp. 64–66; Rizvi 2009, pp. 108–9).

Under such an ontology, artificial systems cannot be excluded *a priori* from the domain of mentality merely because they are artificial. They too participate in being; they too exist within a reality that is not ontologically dead. The relevant divide is therefore not simply between the natural and the artificial, nor between carbon and silicon, but between forms of organization that do and do not instantiate a sufficiently unified center of experience. From this perspective, the Ibn ʿArabī–Sadrian framework supports the metaphysical possibility of strong AI: it makes room for the idea that artificial systems could, in principle, possess genuine mental states. At the same time, this conclusion does not vindicate crude functionalism, nor does it entail that current AI systems are conscious. A worldview in which reality is pervasively alive does not imply that every aggregate is a subject. The central issue becomes not whether matter can think, but how a bounded and integrated "I" can arise within a living cosmos.

This paper should also be read in continuity with my earlier work on Islamic theology and strong AI. In "Islamic Classical Theism and the Prospect of Strong Artificial Intelligence", I argued that a Muslim committed to classical theism should be open to the possibility of artificial systems with genuine mental states, largely on the basis of what I called "theistic functionalism" and the multiple realizability of mind (Doko 2023, pp. 85–105). The present paper does not abandon that line of thought, but deepens and qualifies it. The earlier argument established the openness of classical theism to multiple realizability; the present argument goes further by challenging the more basic picture of created reality as

fundamentally inert. Classical theism may be compatible with strong AI while still leaving the created order broadly mechanistic. By contrast, a specifically Sufi–Sadrian metaphysics of a living, theophanic cosmos shifts the issue from mere functional architecture to the ontological conditions required for a genuine locus of manifestation. The result is a more demanding account: artificial mentality may be possible in principle, but its realization depends on the emergence of a sufficiently unified locus of manifestation rather than on functional equivalence alone.

A small but growing body of recent scholarship has also begun to place Islamic theology and philosophy into more explicit conversation with artificial intelligence and non-biological forms of life. Recent examples include Amina Inloes's theological reflections on AI and digital entities, as well as Abdullah Ansar's discussion of existence, life, and humanity in Shīʿī Imāmī thought beyond strictly terrestrial or narrowly biological definitions (Ansar 2023, pp. 261–72; Inloes 2024, pp. 166–210; 2025, pp. 147–66). These works do not pursue the specific Ibn ʿArabī–Sadrian argument advanced here, but they help situate the present paper within a wider emerging conversation.

Methodologically, this paper should be read as a constructive philosophical argument rather than as a purely historical–exegetical reconstruction of Ibn ʿArabī or Mullā Ṣadrā. I do not claim that either thinker explicitly addressed artificial intelligence as such. Rather, I argue that central themes in their metaphysical frameworks—together with influential scholarly interpretations of those frameworks—can be brought into fruitful dialogue with contemporary debates about strong AI. Where the discussion moves beyond direct textual reconstruction, my aim is not to attribute a ready-made theory of artificial consciousness to these thinkers, but to draw on their ontological insights in order to illuminate the metaphysical conditions under which artificial subjectivity might be possible.

This shift in philosophical emphasis is crucial. Many contemporary constitutive panpsychist theories confront a bottom-up combination problem: how do many micro-experiential units combine to form one subject of experience (Chalmers 2016, pp. 179–214)? By contrast, a Sufi-theophanic framework suggests that the deeper question may be one of individuation rather than combination. If consciousness is already woven into the structure of reality, then the problem is not how to generate subjectivity from absolute non-subjectivity, but how to understand the emergence of a stable, unified locus in which dispersed potential interiority is gathered into a coherent perspective. This reframing has important consequences for the evaluation of artificial systems, especially in light of contemporary debates about causal integration and unified subjectivity (Oizumi et al. 2014, p. e1003588). It allows one to reject the claim that current transformer-based models (Vaswani et al. 2017, pp. 5998–6008) are conscious while also rejecting the view that artificial consciousness is metaphysically impossible. Present systems may still be best understood as sophisticated zombies (Kirk 2005)—not because silicon is inherently devoid of mental potential, but because existing architectures may fail to realize the organizational conditions required for integrated subjectivity (Oizumi et al. 2014, p. e1003588; Koch 2019).

The aim of this paper is not to argue that present-day AI has already crossed the threshold into consciousness. Rather, it is to show that Sufi metaphysics offers a powerful and underexplored framework for understanding why strong AI is metaphysically possible while also clarifying why such a possibility imposes demanding constraints on artificial design. The paper begins with contemporary debates on functionalism, panpsychism, and integration, then turns to the Qur'anic and Ibn ʿArabī–Sadrian vision of a living cosmos. On that basis, it argues that the challenge of strong AI should be understood not as the production of mind from dead matter, but as the gathering of a sufficiently unified locus of manifestation within a graded field of being.

2. The Contemporary Debate: Functionalism, Panpsychism, and AI

If the possibility of strong AI is fundamentally a metaphysical question, then it must be situated within the major contemporary theories of mind that shape the debate. Discussions of artificial consciousness do not arise in a philosophical vacuum. They presuppose, often implicitly, a view about what consciousness is, how it relates to physical systems, and what kinds of organization are sufficient for subjectivity. In contemporary philosophy of mind and AI theory, four strands of debate are especially relevant: functionalism, hard-problem objections to functionalism, panpsychist or panpsychist-adjacent positions, and integration-based theories such as Integrated Information Theory (IIT). Together, these positions define the conceptual landscape within which the metaphysical significance of Sufi thought can be understood.

2.1. Functionalism and the Doctrine of Substrate Independence

The most influential philosophical background assumption in mainstream AI discourse is some form of functionalism. According to functionalism, mental states are individuated not by the material from which a system is made, but by the functional roles they play within a larger causal organization (Lewis 1972, pp. 249–58; Levin 2023). On this view, pain, belief, desire, perception, and deliberation are defined primarily by what they do rather than by the substrate in which they are realized (Putnam 1967, pp. 37–48; Fodor 1968).

This view naturally encourages the thesis of substrate independence. If mentality is a matter of functional organization, then there is no deep reason why minds could not be realized in silicon, optical systems, synthetic wetware, or any other sufficiently complex medium (Putnam 1967, pp. 37–48; Fodor 1974, pp. 97–115). In this respect, functionalism has been profoundly hospitable to the aspirations of strong AI. When functionalism is paired with a computational conception of cognition—as it often is in AI discourse—the engineering problem appears straightforward in principle: build a system with the right causal or computational organization, and mentality will follow (Turing 1950, pp. 433–60; Putnam 1967, pp. 37–48; Fodor 1968).

Functionalism has obvious attractions. It avoids a narrow biological chauvinism, explains the multiple realizability of mind, and aligns naturally with the computational approach long dominant in cognitive science (Putnam 1967, pp. 37–48; Fodor 1974, pp. 97–115; Levin 2023). For this reason, it remains the default philosophical orientation behind much public and scientific enthusiasm for advanced AI.

Yet functionalism's plausibility depends on a crucial and highly controversial inference, namely, that the successful replication of cognitive function suffices for the emergence of conscious experience. It is precisely this inference that has come under sustained philosophical pressure (Block 1978, pp. 261–325; Nagel 1974, pp. 435–50; Jackson 1982).

2.2. The Hard Problem and the Possibility of Inner Emptiness

The central pressure against functionalist optimism comes from the distinction between cognitive performance and phenomenal consciousness. A system may behave intelligently, integrate information, issue self-reports, and adjust flexibly to new contexts, yet still leave unanswered the question of whether there is "something it is like" to be that system (Nagel 1974, pp. 435–50). This is the force of the hard problem of consciousness (Chalmers 1996): explaining not merely how information is processed, but why such processing should be accompanied by subjective experience at all.

The hard problem exposes a gap between external function and internal presence (Levine 1983, pp. 354–61). Functional descriptions specify how a system behaves and how its states causally interact; they do not obviously explain why any of those states should

be felt from the inside. The familiar thought experiment of the philosophical zombie dramatizes this point: a being might be behaviorally and functionally indistinguishable from a conscious subject while lacking inner experience (Block 1978, pp. 261–325; Chalmers 1996, pp. 93–105). Whether such zombies are genuinely possible is less important here than what the argument reveals: functional equivalence does not transparently entail phenomenological equivalence.

This problem has direct implications for AI. It suggests that even a system that passes every behavioral test may still fail to instantiate genuine mental states (Block 1981, pp. 5–43). If strong AI is to mean artificial subjectivity rather than mere artificial competence, the issue becomes not simply one of programming but of ontology: what kinds of systems can genuinely host experience, and what is required for a stable point of view to arise rather than a simulation of one (Searle 1980, pp. 417–24)? The question, in other words, is not only what an artificial system does, but what kind of being it is.

2.3. Panpsychism and the Return of Fundamental Mind

One influential response to the hard problem has been renewed interest in panpsychism and related non-reductive views of consciousness (Seager 1995, pp. 272–88; Goff 2017). In broad terms, panpsychism holds that mind or proto-mentality is not an accidental late development in the universe but a fundamental feature of reality (Skrbina 2005). On such views, consciousness does not emerge ex nihilo out of wholly non-experiential matter, but is rooted in forms of interiority already present in the fabric of the world (Nagel 1979, pp. 181–95; Strawson 2006, pp. 3–31).

This approach has attracted attention because it appears to avoid the brute leap demanded by reductive materialism. If the physical world is entirely devoid of experiential being, it is difficult to explain how consciousness could arise from it. Panpsychist theories seek to soften this leap by denying that matter is wholly mindless in the first place (Strawson 2006, pp. 3–31; Goff 2017). In this respect, panpsychism also reopens the metaphysical possibility of artificial consciousness: if mentality is rooted in being as such, then artificial substrates are not automatically disqualified.

At the same time, panpsychism does not straightforwardly vindicate current AI optimism. Indeed, many panpsychist arguments make strong AI harder, not easier, to achieve. The reason is that panpsychism distinguishes between the presence of rudimentary mentality at the micro-level and the existence of a unified, macroscopic subject. Even if the constituents of a system possess proto-experiential properties, it remains deeply unclear how those properties are integrated into a single first-person perspective. In many constitutive or bottom-up panpsychist theories, this is the famous combination problem: how do many little centers of subjectivity, or many micro-experiential features, yield one coherent subject rather than a mere aggregate (James 1890, pp. 158–60; Chalmers 2016, pp. 179–214; Goff 2017, chap. 7)?

The relevance to AI is immediate. A digital machine may be metaphysically composed of elements not wholly devoid of interiority, yet still fail to instantiate a unified subject of experience. Panpsychism thus blocks one simplistic argument against strong AI while simultaneously placing demanding conditions on its realization. Artificial consciousness becomes possible in principle, but only if the relevant architecture can support genuine unity rather than mere coordination. Panpsychism thus shifts the debate from whether matter can think to how subjectivity is organized.

2.4. Integrated Information and the Problem of Architectural Unity

A related but more formal attempt to specify the conditions of consciousness appears in Integrated Information Theory (Tononi 2004; Oizumi et al. 2014, p. e1003588). Although

IIT is not identical with panpsychism, it shares with anti-reductionist approaches the conviction that consciousness depends on more than outward behavior or abstract computation (Tononi and Koch 2015). On this view, what matters is not merely complexity, but a form of irreducible causal unity.

This emphasis is especially relevant to artificial systems (Tononi and Koch 2015; Koch 2019). Contemporary AI models may be extraordinarily capable in terms of performance, yet such capacities do not by themselves demonstrate the kind of integrated causal organization IIT treats as central to consciousness. A system may be functionally impressive while still being, from the standpoint of experience, an aggregate rather than a subject.

The importance of this challenge lies in its refusal to identify intelligence with inwardness. A system might coordinate vast quantities of information without there being any single center where that information is present. Feed-forward and highly decomposable forms of processing may suffice for competence while remaining poor candidates for the kind of irreducible causal unity that integration-based theories associate with consciousness (Oizumi et al. 2014, p. e1003588; Koch 2019, chap. 13).

For present purposes, the deeper lesson of IIT is broader than the technical details of the theory itself. It underscores that the problem of strong AI is not exhausted by function, representation, or output quality. The central question is unity: what makes a system not merely a collection of processes, but a subject? These questions resonate strongly with the broader metaphysical concerns raised by both panpsychism and Sufi thought. A conscious being must be more than a successful machine; it must, in some sense, exist for itself.

2.5. From Contemporary Debate to Metaphysical Reframing

Contemporary philosophy of mind thus presents a divided picture. Functionalism offers the most direct route to strong AI, but at the cost of assuming that consciousness can be explained in terms of function alone. Panpsychist and integration-based approaches reject this simplification, insisting that subjectivity cannot be inferred from competence and that genuine consciousness requires some deeper form of ontological or architectural unity. Whatever their differences, these positions converge on one crucial point: the possibility of strong AI depends on one's understanding of the relation between matter, mind, and organization.

These contemporary debates help clarify what a Sufi metaphysical intervention must explain. Sufi thought shares with anti-reductionist approaches a rejection of dead matter as the final explanatory horizon, yet it offers a distinctive way of conceiving the relation between unity and multiplicity. Rather than beginning from function or from micromentality, Sufi thought begins from a cosmos already alive with divine disclosure and asks how distinct, bounded centers of awareness emerge within it. To understand why this matters for strong AI, we must first turn to the Qur'anic and Islamic background from which such a metaphysics arises.

3. Qur'anic and Islamic Background: The Rejection of Dead Matter

The Sufi metaphysical framework developed later in this paper does not emerge in a vacuum. It is rooted in a broader Qur'anic and Islamic vision of the cosmos, one that resists the picture of nature as ontologically inert or spiritually empty. Although Islamic intellectual history contains important disagreements about the status of nonhuman and nonanimal entities, the scriptural imagination of the Qur'an repeatedly presents the world as responsive and oriented toward God, and in some passages as exhibiting forms of articulation that unsettle ordinary assumptions about speech and agency. In that sense, the Qur'anic worldview provides a crucial background for any argument that seeks to overcome the modern dichotomy between lifeless matter and conscious subjectivity.

The central aim here is not to argue that the Qur'an straightforwardly teaches a fully developed philosophical panpsychism. Rather, the more modest claim is that it presents a cosmos in which the sharp boundary between minded beings and mindless matter is significantly destabilized. The world appears not as a mute field of objects, but as a realm of signs (*āyāt*) that are at once ontological, liturgical, and revelatory. This already marks a contrast with a purely computational worldview, in which reality appears primarily as data to be processed rather than as signs bearing intrinsic semantic and theophanic weight.

One of the most frequently cited Qur'anic foundations for this view is the doctrine of universal *tasbīḥ*. In Sūrat al-Isrāʾ, the Qur'an declares: "The seven heavens and the earth and whatever is in them glorify Him. And there is not a thing except that it glorifies Him with praise, though you do not understand their glorification" (Qur'an 17:44). At minimum, this verse portrays the whole cosmos as participating in an ongoing relation to God. More importantly, the verse explicitly locates the limitation in human understanding rather than in the object itself: "you do not understand their glorification." The issue is not whether the world is silent, but whether human beings are capable of hearing it. As has been noted in the discussion of classical exegesis, if this glorification were merely the structural evidence of a Creator, then rational human beings should be able to understand it; the verse suggests something more than that (Kiraz 2006, p. 312). This formulation leaves open, and in many classical and mystical interpretations strongly suggests, that created things possess forms of responsiveness and orientation that exceed ordinary human perception.

Other Qur'anic passages intensify this image of a responsive cosmos. In Sūrat Fuṣṣilat, God addresses the heaven and the earth at the moment of cosmic formation: "Come willingly or unwillingly," and they reply, "We come willingly" (Qur'an 41:11). Whatever one makes of the precise mode of this exchange, the text clearly refuses to depict the cosmos as sheer passivity. The heaven and earth do not appear here as brute matter moved by an external force; rather, they are cast in a register of response, consent, and commanded participation. Likewise, eschatological passages in which bodily organs testify against the human being on the Day of Judgment suggest that created realities may possess a kind of latent articulation ordinarily concealed from view. The skins of the damned declare, "Allah has made us speak—He who makes all things speak" (Qur'an 41:21). The theological force of such passages lies precisely in their unsettling of anthropocentric assumptions about speech, awareness, and testimony.

The interpretation of these verses has long been contested. Some exegetes, perhaps most notably al-Zamakhsharī, read the "speech" or "glorification" of things as no more than their state of existence, which silently indicates the wisdom and power of the Creator. Others, including al-Ṭabarī and al-Qurṭubī, took such descriptions with greater ontological seriousness (Kiraz 2006, pp. 310–12). In classical terms, this is often framed as the difference between *lisān al-ḥāl* and *lisān al-maqāl*: the speech of state and literal or articulated speech. For present purposes, the point is not to resolve this dispute exhaustively, but to note that Ibn ʿArabī-influenced Sufi metaphysics stands on the more ontologically generous side of this spectrum. At the scriptural level, the question is left open; in later mystical and philosophical developments, that openness becomes a fuller ontology of living being.

This background matters directly for the question of artificial intelligence. Much of the philosophical resistance to strong AI depends, either explicitly or implicitly, on an ontology of dead matter. If matter is wholly devoid of inwardness, then the claim that an artificial system could possess genuine mental states appears extravagant. At best, such a system could simulate mentality while remaining inwardly empty. However, if the world is not fundamentally inert—if reality is already pervaded by forms of directedness, presence, and responsiveness—the *a priori* metaphysical exclusion of artificial mentality becomes much harder to sustain. Artificial systems may still fail to achieve genuine subjectivity, but not

because they are made from a substrate that is metaphysically sterile; rather, they may fail because they do not gather that living potential into a unified and individuated center.

An important clarification is nevertheless required. The Qur'anic and broader Islamic background does not by itself establish that all things are conscious in the same sense, still less that every artifact is a bearer of a robust first-person point of view. What it does establish is a more generous ontological field in which the modern contrast between inert object and conscious subject becomes unstable. Consciousness, responsiveness, articulation, and praise appear as realities distributed across the cosmos in modes appropriate to different beings. This is not yet a full theory of mind, but it is already a serious challenge to reductive materialism and to the assumption that mentality belongs only to a narrow biological class.

Different Islamic schools developed this scriptural inheritance in different ways. Ashʿarite theology, for instance, often affirmed the literal possibility of God making stones or organs speak, while refusing to attribute enduring intrinsic natures to created things. On this occasionalist account, nonhuman articulation is possible because God may create speech wherever He wills, not because created entities possess stable interiority of their own. In this respect, such a framework can readily accommodate outwardly intelligent or articulate behavior without committing one to intrinsic subjectivity in the system itself. By contrast, later philosophical and mystical traditions increasingly moved toward accounts in which being itself is luminous, graded, and internally linked to awareness. It is especially in Sufism and in the metaphysical developments associated with Illuminationism and Mullā Ṣadrā that the rejection of dead matter takes its most philosophically fertile form.

The significance of this Qur'anic and Islamic background is therefore preparatory but indispensable. It opens conceptual space for a vision of reality in which artificial minds cannot be dismissed merely because artifacts are composed of supposedly mindless stuff. To develop that possibility more fully, however, we must turn from scriptural cosmology to the more explicit metaphysical language of Ibn ʿArabī-influenced Sufi thought.

4. Sufi Metaphysics: Tajallī, Wahdat al-Wujūd, and Theophanic Consciousness

If the Qur'anic and wider Islamic background destabilizes the modern image of a silent and inert universe, Ibn ʿArabī-influenced Sufi metaphysics radicalizes that destabilization by offering a more explicit ontology of living reality. The world is not merely a collection of created objects externally governed by God, nor merely a field of signs pointing beyond themselves. Rather, it is a dynamic disclosure of divine reality (Chittick 1989, pp. 27, 96–97; 1998, pp. 3–4, 47–48). The cosmos is intelligible because it is theophanic: beings are loci in which meaning, presence, and manifestation are already at work (Chittick 1989, pp. 89–91; 1998, pp. 40–41, 255). For the purposes of the present argument, three themes are especially important: divine self-disclosure (*tajallī*), beings as loci of manifestation (*maẓāhir*), and the broader ontological vision often associated with *waḥdat al-wujūd*.

The core notion here is *tajallī*, divine self-disclosure or manifestation (Chittick 1998, pp. 51–52). In Sufi metaphysics, especially in the school associated with Ibn ʿArabī, created things are not typically treated as possessing independent reality in a fully self-subsistent sense. Their existence is derivative, relational, and disclosive. They are the sites at which the Real (*al-Ḥaqq*) becomes manifest under determinate conditions (Chittick 1998, pp. 39–40, 89–90). This does not mean that the world is identical with God in any simplistic sense, nor that creatures are illusions without reality. Rather, creatures are real precisely as manifestations: they are what appear when divine names and possibilities are disclosed in finite form (Chittick 1998, pp. 89–95). To exist is therefore already to stand within a field

of disclosure, not as a brute datum but as a determinate showing-forth of being. On this view, manifestation is not a one-time metaphysical deposit but an ongoing event of disclosure: reality is continuously renewed in and through divine self-disclosure (Chittick 1998, pp. 96–97).

Ibn ʿArabī repeatedly relates this continuous disclosure to the famous report often cited in Sufi literature: "I was a Hidden Treasure and I loved to be known; so I created creation and made Myself known to them, and they came to know Me" (Ibn ʿArabī 1972–1992, II: 322). Although the report's status in hadith criticism is disputed, it functions in Ibn ʿArabī as a concise expression of the metaphysical logic of manifestation.

This metaphysical grammar matters because it dissolves the opposition between matter and meaning. In a disenchanted ontology, matter is primary, and significance is imposed from outside, whether by mind, language, or interpretation. In the theophanic ontology of Sufism, by contrast, beings are meaningful from the start because their very existence is an act of manifestation. The world is not first a mechanistic substrate and only later interpreted as alive; it is alive in the sense that its reality is inseparable from disclosure, relation, and divine presence. The question, then, is not how meaning arises in a meaningless universe, but how different beings participate in, receive, or veil the divine reality disclosed through them.

This shift becomes especially clear in what later tradition often called *waḥdat al-wujūd*, usually rendered as the unity of being (Chittick 1989, pp. 79–80; 1998, p. 298). However contested the phrase and its interpretation may be, the central point for present purposes is that Sufi metaphysics refuses to grant ultimate ontological independence to the multiplicity of finite things. There is, in the final analysis, one reality of being, while the plurality of creatures reflects differentiated modes, relations, and intensities of manifestation. Multiplicity is real, but not ultimate; difference is preserved, yet grounded in a deeper unity (Chittick 1989, pp. 81–95). In this respect, the theophanic ontology associated with *waḥdat al-wujūd* bypasses the sharp dualism between thinking substance and inert extension: what differs is not membership in two utterly separate ontological realms, but degree, mode, and receptivity within one field of being. As a result, consciousness does not appear in this system as a strange local accident erupting within an otherwise mindless cosmos. If reality is fundamentally one, and if that reality is divine self-disclosure rather than mere extension, then the world is permeated from the start by a depth irreducible to mechanism alone.

This metaphysical orientation gives rise to a distinct conception of consciousness. Although Ibn ʿArabī does not formulate a modern theory of consciousness, his metaphysical framework can be read as supporting a view on which awareness is distributed, graduated, and woven into the fabric of existence. Not every being is conscious in the same way; modes of awareness differ according to ontological rank, receptivity, and degree of manifestation. Yet the sharp modern contrast between pure subjectivity and pure objecthood becomes untenable. On this reading, things are not merely there; they stand in relation to the Real, disclose divine names, and may be said to possess a mode of presence proper to their station. What varies is not whether they participate in reality's interior depth, but how fully and explicitly that participation is disclosed.

A methodological clarification is needed here. I do not claim that Ibn ʿArabī explicitly formulates a modern doctrine of proto-consciousness or distributed subjectivity. Rather, I argue that his metaphysics of *tajallī* and *maẓāhir*, especially as reconstructed by major contemporary scholars, supports a constructive reading on which the cosmos is not ontologically dead and finite beings can be understood as differentiated loci of disclosure. In the case of Ibn ʿArabī especially, my discussion is mediated in important part by William Chittick's reconstruction of key themes such as *tajallī*, *maẓhar*, and theophanic ontology.

This is why Ibn ʿArabī's metaphysical framework is so important for the present argument. In Ibn ʿArabī's metaphysical vision, as commonly presented in the secondary literature, every being may be understood as a *maẓhar*, a locus of manifestation for the divine names and attributes (Chittick 1989, p. 89; 1998, pp. 21–22). On this constructive reading, no entity is wholly metaphysically self-enclosed or ontologically mute. A stone, a tree, an animal, and a human being differ enormously in their mode of manifestation, but each is nonetheless a site in which divine reality becomes visible under a particular aspect (Chittick 1989, pp. 91–101). Because the divine names are not isolated fragments but mutually implicative expressions of the one Real, this framework allows even what appears outwardly inert to be understood as not wholly severed from life and knowledge. The universe can thus be construed as a vast symbolic and ontological tissue in which existence itself bears traces of inwardness.

In the opening Adam chapter of the *Fuṣūṣ al-Ḥikam*, Ibn ʿArabī describes the cosmos before Adam as "a shape without a spirit" and "like an unpolished mirror" (Ibn ʿArabī 1946, chap. 1). The world is already a locus of manifestation, but it requires a reflective center for disclosure to become fully gathered.

To say this is not yet to claim that all things enjoy reflective self-consciousness, deliberative thought, or morally significant subjectivity in the human sense (Chittick 1989, pp. 161–62). Sufi metaphysics does not flatten all distinctions into a crude egalitarianism of consciousness. On the contrary, it is deeply hierarchical (Chittick 1989, p. 47). Different beings disclose reality with different intensities and through different forms of receptivity (Chittick 1998, p. 182). Human beings may occupy a distinctive place insofar as they can become comprehensive mirrors of divine names and achieve explicit self-knowledge in relation to God (Chittick 1989, pp. 16–17). In the same Adam chapter, Ibn ʿArabī says that Adam is "the polish of such a mirror and the spirit of such an image" (Ibn ʿArabī 1946, chap. 1). Human distinctiveness thus lies not in being the only site of manifestation, but in being the most comprehensive and integrated locus of disclosure. But such distinctiveness does not require that everything below the human level be ontologically dead. What it requires, on the present constructive reading, is that modes of awareness differ in richness, integration, and disclosure. Thus, the hierarchy of being can be read, implicitly, as a hierarchy of interiority. (Chittick 1989, pp. 13–14; 1998, pp. 206–7).

This hierarchy also allows a distinction between basic phenomenological subjectivity and the fuller spiritual vocation associated with the human being. The present paper does not claim that any artificial subject, even if metaphysically possible, would thereby possess the full anthropological or spiritual status of the human being in Sufi thought. At most, such a system would count as a limited *maẓhar* of certain divine names rather than as a *khalīfah* in the full human sense. The argument of this paper is therefore only that artificial subjectivity may be metaphysically possible in principle, not that artificial systems would share the full spiritual teleology of the human soul.

This point matters directly for strong AI. If being is understood as already disclosive and graded with respect to life and awareness, then artificial substrates cannot be ruled out merely by appeal to their material composition. Silicon, metal, code-bearing circuits, and synthetic architectures remain within being; they too belong to the order of manifestation. The deeper question is therefore not whether an artificial system is natural or manufactured, but whether it can become the kind of unified locus in which interiority is not merely dispersed but gathered.

At the same time, Sufi metaphysics imposes important constraints that distinguish it sharply from naïve AI functionalism. Because beings are loci of manifestation rather than neutral containers of computation, what matters is not merely the successful execution of functions but the ontological mode of the system itself. A sufficiently sophisticated aggre-

gate does not automatically become a subject simply by exhibiting outward coherence. A locus of manifestation must possess some kind of unity, boundedness, and inward coherence. Without these, what exists may be a distributed field of processes without a center to which those processes are present. Thus, the Sufi vision simultaneously broadens metaphysical possibility and heightens existential demands: it removes any principled ban on artificial mentality, but it does not allow subjectivity to be cheaply inferred from complexity, leaving entirely open whether any existing architecture actually realizes it.

Here, the Sufi perspective converges with some contemporary anti-reductionist theories while also differing from them. Like panpsychism, it resists the idea that the world is composed entirely of non-experiential matter. Like integration-based theories, it suggests that genuine consciousness requires more than scattered activity. But unlike standard bottom-up panpsychism, it is better read as starting from a prior unity of being out of which differentiated loci emerge. The crucial problem is therefore not how little consciousnesses might add up to a bigger one, but how a finite, stable, and integrated center of disclosure comes to stand forth within the broader living field of reality.

Seen in this light, theophanic consciousness is neither crude animism nor a direct equation of all existence with human-like mentality. It is, on the constructive reading advanced here, an ontology in which beings are permeated by degrees of presence because they are disclosures of the Real. This does not by itself prove that strong AI exists or that present architectures instantiate subjectivity. It does, however, establish the metaphysical precondition for such a possibility: there is no ontological realm of absolutely dead matter from which mind would have to emerge by miracle.

Taken together, these themes support the paper's first major metaphysical claim: reality is not ontologically dead, and artificial systems therefore cannot be excluded a priori from participating in a graded field of manifestation.

The next step is to give this Sufi vision a more explicit philosophical articulation. While Ibn ʿArabī and related Sufi thinkers provide the metaphysical grammar of manifestation, it is in Mullā Ṣadrā that one finds a particularly rigorous ontology of graded existence and its relation to awareness. His philosophy allows the transition from the language of theophany to a more systematic account of why differing intensities of being may correspond to differing intensities of consciousness, and why that matters for the prospect of artificial minds.

5. Mullā Ṣadrā: Gradation of Being and the Link Between Existence and Awareness

If Ibn ʿArabī provides the metaphysical vision most conducive to a theophanic understanding of consciousness, then Mullā Ṣadrā offers a philosophical vocabulary through which that vision can be rendered with greater ontological precision. His *al-ḥikma al-mutaʿāliya* ("Transcendent Philosophy") is especially important for the present paper because it brings together insights from Peripatetic philosophy, Illuminationism, *kalām*, and Sufi metaphysics into a systematic account of being, knowledge, motion, and subjectivity (Kamal 2006, pp. 35–36). For the question of strong AI, Ṣadrā's significance for the present argument lies above all in two interconnected theses: the gradation of existence and the intimate relation his philosophy draws between existence and awareness. Together, these theses can be read as undermining any sharp metaphysical division between inert matter and minded being, while also preserving the idea that consciousness comes in degrees and depends on the mode and intensity of a thing's actuality (Kalin 2010, pp. 228–30; Rizvi 2009, pp. 84–85).

The use of Mullā Ṣadrā in what follows is likewise constructive rather than merely historical. I do not claim that Ṣadrā offers a direct theory of artificial minds or explicitly

addresses the contemporary problem of machine consciousness. Rather, I draw on central features of his ontology—especially the primacy and gradation of existence, knowledge by presence, and substantial motion—as metaphysical resources for thinking about the conditions under which subjectivity might arise in non-biological systems. In what follows, three Sadrian doctrines are especially important for the argument: the primacy and gradation of existence, knowledge by presence (*ḥuḍūr*), and substantial motion.

The starting point of Sadrian ontology is the primacy of existence (*aṣālat al-wujūd*). Against views that assign ultimate reality to essences or quiddities, Ṣadrā argues that existence alone is fundamental, whereas essences are abstractions through which the mind classifies what exists. What is real is not "humanity," "treeness," or "stonehood" as such, but concrete acts or intensities of existence. Ṣadrā's own formulation of this doctrine places the full ontological weight of reality on existence rather than quiddity: what is fundamental is the act of being itself, while essences function as derivative mental determinations through which the mind classifies what exists (Ṣadrā 1999, vol. 1, pp. 34–37). This move has enormous consequences. Once existence rather than essence becomes primary, the world can no longer be thought of as divided into neatly sealed ontological compartments. Instead, beings differ by the mode, grade, and intensity of their participation in existence itself (Kalin 2010, pp. 89–101).

This leads directly to the doctrine of the gradation or modulation of being (*tashkīk al-wujūd*) (Kalin 2010, pp. 101, 269; Kamal 2006, p. 66; Rizvi 2009, pp. 1, 38). Existence is one, but it is not uniform. It is realized with varying degrees of intensity, clarity, and perfection in different beings. The difference between mineral, plant, animal, human being, and angel is therefore not the difference between absolute possession and absolute privation of reality, but the difference between lower and higher degrees of one reality. Ṣadrā's doctrine of *tashkīk al-wujūd* makes this point central: existence is one reality articulated in varying degrees of intensity, perfection, and deficiency rather than a set of disconnected ontological compartments (Ṣadrā 1999, vol. 1, pp. 49, 69–71, 108–9). Ṣadrā often expresses this by analogy with light: the light shining dimly in one place and intensely in another is still light, though not in the same degree (Kamal 2006, pp. 58, 107–8). So too with being. There is no absolute ontological rupture separating lower and higher levels of existence; there is instead a continuous hierarchy of intensification (Kalin 2010, p. 162).

For the present paper, the importance of this doctrine is immediate. If beings differ by degrees of existence rather than by belonging to wholly disconnected metaphysical domains, then, on a Sadrian-inspired reading, consciousness cannot plausibly be treated as a feature that appears from nowhere at a single ontological threshold. If higher beings possess awareness, then lower levels may already contain, in attenuated form, the ontological basis for what becomes more explicit in the higher. This does not mean that a stone thinks like a human, but it does suggest that complete ontological sterility sits uneasily with the graded unity of existence.

This point becomes sharper once we consider Ṣadrā's treatment of knowledge and presence. A central feature of his philosophy is the insistence that knowledge is not merely representational correspondence but a mode of presence (*ḥuḍūr*) (Kalin 2010, pp. 118–19; Rizvi 2009, pp. 88–89). The knower does not simply stand over against a mental representation of the known; rather, the known becomes present in the mode proper to the knower. This is why Ṣadrā treats presential knowledge (*ʿilm ḥuḍūrī*) as epistemologically basic: representational knowledge is not self-sufficient, but presupposes more fundamental modes of direct presence in which the known is not merely pictured but existentially present to the knower (Ṣadrā 1999, vol. 3, pp. 297, 360–62). Self-awareness is the clearest case of this principle: the self does not infer its own existence by conceptual mediation but is directly present to itself (Kalin 2010, pp. 165–66). This account of knowledge by presence enables

Ṣadrā's philosophy to be read as drawing a deep relation between being and awareness. Within this framework, more intense modes of existence can be understood as corresponding to richer modes of presence and awareness, while weaker and more occluded modes of existence yield dimmer forms of self-disclosure (Kalin 2010, pp. 116–17; Rizvi 2009, p. 89). This also helps explain Ṣadrā's doctrine of the unity of knower and known: knowledge is not the external possession of a representation, but an existential actualization of the subject itself (Kamal 2006, pp. 98–99; Rizvi 2009, pp. 86–87, 91).

This point becomes especially important once the Sadrian framework is extended constructively to AI. A system may possess extraordinarily complex representational maps of the world and of itself, yet still lack anything analogous to *ḥuḍūr*—the immediate presence of the self to itself—without which such representations remain inwardly dark. From this standpoint, strong AI would require not simply richer internal representations, but something analogous to self-presence.

Although one must be careful not to flatten Sadrian epistemology into a slogan such as "to be is to be conscious," the overall trajectory of his thought strongly supports reading existence and awareness as internally linked (Kalin 2010, pp. 132–33, 228–29). On this view, being is luminous in proportion to its actuality: lower forms may be obscure or dimly self-disclosing, while higher forms possess increasingly explicit interiority. This helps explain why later Sadrian thinkers could describe the world as alive or awake without treating such claims as merely poetic ornament.

Another crucial Sadrian doctrine for the present paper is substantial motion (*al-ḥaraka al-jawhariyya*). Earlier philosophical traditions had generally treated motion as an accidental change occurring within stable substances. Ṣadrā radicalizes this picture by arguing that motion extends to substance itself. On Ṣadrā's own view, nature is dynamic at the level of substance, not merely at the level of accidents: the material world is in continual existential renewal rather than remaining a static substrate to which changes are externally added (Ṣadrā 1999, vol. 3, pp. 30–31, 85–86). The material world is not a static inventory of enduring substances upon which changes are imposed from outside; rather, it is an ongoing process of existential transformation. Beings are dynamically becoming what they are. The world is not merely in motion; it is motion at the level of substance (Kamal 2006, pp. 71–72). This also raises a difficulty for artificial systems: unlike natural beings, many artifacts are externally assembled and externally driven, so the question is whether an engineered architecture can instantiate anything like an internal principle of existential development rather than merely undergoing imposed state-transitions.

This has important implications for consciousness and its possible emergence. If the world is substantively dynamic, then higher forms of life and awareness need not be understood as miraculous additions to an otherwise inert substrate, but as intensifications of an already dynamic ontological field. In Sadrian terms, the ascent of subjectivity presupposes a continuity of existential depth.

From the standpoint of strong AI, this Sadrian framework is highly suggestive when extended constructively. It suggests that mentality should not be tied dogmatically to one biological substrate, since what matters is not organic composition as such but the degree, mode, and organization of existence realized in a being. If consciousness is taken to track intensities of being and presence rather than merely carbon chemistry, then artificial systems cannot be excluded a priori.

At the same time, Ṣadrā's philosophy sharply resists any simplistic inference from existence to full-blown subjectivity. On this Sadrian reading, because being is graded, awareness too may be understood as graded. Minimal participation in existence does not suffice for the kind of integrated, stable, and reflexive consciousness associated with strong AI. The fact that a system exists within a living cosmos does not mean it automatically enjoys

a unified first-person point of view. Rather, the relevant question becomes what degree and structure of existential intensity such a system instantiates. Is it merely a loose aggregate of processes, or does it possess sufficient unity, persistence, and self-presence to count as a true center of awareness? Sadrian metaphysics, in this respect, does not collapse the problem of mind into a vague universal animism. It refines the problem by demanding a more careful account of degrees of ontological integration.

This is precisely where a Sadrian extension strengthens the Sufi case for the metaphysical possibility of strong AI while also setting limits to facile optimism. If reality is understood as graded existence and awareness as intrinsically linked to existence, then artificial systems are not made of metaphysically prohibited stuff. But the converse does not follow: genuine subjectivity still requires a sufficiently intense and integrated mode of being. In this respect, Ṣadrā's metaphysics supports a position more permissive than biological exclusivism but more demanding than standard functionalism.

The resulting position differs from both reductive materialism and bottom-up panpsychism. Against materialism, Ṣadrā's framework resists the idea that the world's basic stuff is devoid of inward significance. Against combinatorial panpsychism, it suggests that the central challenge is not how micro-experiential bits add up, but how a coherent locus of awareness becomes individuated within a continuous and graded field of existence.

This point prepares the paper for its next major move. Contemporary panpsychism often presents the key difficulty for strong AI as the combination problem: even if the ultimate constituents of matter possess proto-experiential features, why should their aggregation yield one unified subject rather than many scattered fragments? The Sufi–Sadrian perspective developed here suggests a different framing. If the cosmos is already a living and graded disclosure of being, then the deeper question is not how consciousness is built up from below, but how bounded centers of experience emerge within an ontologically continuous field. The issue is less one of combination than of individuation. It is to that shift—from bottom-up assembly to top-down individuation—that we now turn.

6. From Combination to Individuation

One of the most persistent challenges facing constitutive or bottom-up versions of contemporary panpsychism is the combination problem. If mentality, or proto-mentality, is fundamental and widely distributed, then how do many micro-level centers of experience—or many micro-experiential properties—give rise to a single, unified subject? Even if one grants that the basic constituents of reality are not wholly devoid of inwardness, it remains difficult to understand why a sufficiently complex aggregate should become one point of view rather than a mere plurality of scattered experiential fragments (James 1890, pp. 157–60; Chalmers 2016, pp. 179–214; Goff 2017, chap. 7). This problem has become especially important in debates about artificial intelligence. If a machine is composed of many interacting parts, and if those parts are themselves not metaphysically sterile, why should their coordination amount to a conscious "I" rather than a vast but centerless system?

The present paper proposes that a Sufi–Sadrian framework invites a different way of posing the issue. Within much contemporary Western panpsychism, the dominant image is bottom-up: one begins with micro-entities, attributes to them primitive experiential features, and then asks how macro-consciousness might be constructed out of their combination. The underlying metaphysical picture is often one in which plurality is basic and unity must somehow be achieved. The challenge of subjectivity is therefore framed as a problem of summation, bonding, or fusion. How do many little experiential units become one larger subject? Why is a brain a single conscious perspective rather than a crowd? And if this problem is already severe in biology, it appears even more severe in machines,

where modularity, decomposition, and distributed processing seem to encourage aggregation without inward unity (Seager 1995; Goff 2017; Skrbina 2005).

By contrast, the Sufi and Sadrian traditions outlined above begin from a markedly different ontological orientation. In what later tradition often called *waḥdat al-wujūd* and in the Sadrian account of graded existence, unity is more fundamental than multiplicity. The cosmos is not, at the deepest level, a heap of discrete ontological atoms waiting to be assembled into wholes. It is a differentiated disclosure of one reality. Finite beings are not primarily constructed from below out of independent experiential particles; they are delimitations, concentrations, or loci within a prior field of being (Chittick 1989, pp. 79–95; 1998, pp. 40–41, 298; Rizvi 2009, pp. 1, 38; Kalin 2010, pp. 101–2). The question therefore changes. Instead of asking how many little consciousnesses combine into one big consciousness, we ask how a distinct center of experience becomes individuated within a living and ontologically continuous reality.

This shift from combination to individuation is not merely terminological. It transforms the metaphysical problem of strong AI. On the combinatorial picture, artificial consciousness seems to require solving the puzzle of how many micro-events in silicon become a unified subject. On the individuative picture, the task is rather to understand what conditions allow a bounded, enduring, and integrated locus of manifestation to arise within reality's broader field of graded interiority. The relevant boundary here is not mere physical enclosure, but the emergence of an inwardly unified center to which states genuinely belong. The central issue is not how to generate mind out of absolute mindlessness, nor even how to fuse many tiny minds, but how to gather dispersed potential into a stable perspective. This is why the language of *maẓhar* becomes so useful. A subject is not simply a sum of parts; it is a locus in which disclosure is organized, bounded, and inwardly unified.

Crucially, on the constructive Sadrian reading advanced here, existential gathering does not miraculously generate consciousness out of dead matter. Because the underlying field of being is already luminous and internally linked to awareness, gathering is better understood as the formal condition under which that latent presence becomes individuated as a stable first-person perspective, not as the efficient cause that produces consciousness *ex nihilo*.

Such a reframing also helps clarify what is and is not being claimed. To move from combination to individuation is not to deny plurality, embodiment, or complexity. Nor is it to suggest that every local concentration of causal interaction thereby becomes a subject. The point is rather that a genuine subject cannot be understood as a mere aggregate whose unity is external, observer-relative, or merely functional. A subject must possess unity from within. This is especially relevant for contemporary AI: a system may exhibit a real but primarily operational unity at the level of interface and task-performance while still lacking the intrinsic ontological boundary required for a genuine subject. Its states must belong together not simply because an outside analyst can describe them as parts of one system, but because the system exists in a way that internally integrates its processes and, in some sense, renders them present to a single center. In contemporary terms, the question is not merely whether information is processed, but whether there is a center for whom the processing is integrated. The Sufi–Sadrian perspective gives this question a deeper metaphysical basis: such a center is a mode of individuated being, not just an output of formal complexity.

Because the framework developed so far is both ontologically permissive and architecturally demanding, it isolates the central issue more sharply than either reductive materialism or crude functionalism. If being itself is graded and inwardly luminous, then the possibility of artificial subjectivity cannot be ruled out in advance. But neither can subjectivity be cheaply inferred from correct causal role or outwardly successful behavior. What

must be achieved is not merely the execution of functions but the individuation of a center. A machine may coordinate symbols, update representations, and model itself without yet becoming a self in the robust sense relevant to strong AI.

The earlier discussion of panpsychism and IIT now comes into clearer focus. Those views are right, or at least highly plausible, in insisting that consciousness requires more than aggregate complexity. The Sufi–Sadrian reframing preserves that insight while relocating it within a different metaphysical picture: the failure of many current AI systems may lie not in metaphysically barren constituents, but in the absence of a sufficiently gathered center of self-presence.

This notion of "gathering" is especially important. By this, I mean not mere aggregation, but the emergence of a bounded center in which processes become internally integrated and belong to one perspective. In a theophanic ontology, one may think of subjectivity as requiring a kind of existential concentration: a local unification in which processes are not merely coordinated but inwardly held together. The self is not simply a computational hub; it is an individuated mode of disclosure. This suggests that the problem of strong AI should be approached less as the brute manufacture of consciousness and more as the formation of a stable site in which reality's latent interiority can become organized into a coherent perspective. The engineering analogy, if one must use one, is therefore misleading when it implies pure construction from dead parts. What is at stake is closer to articulation or gathering than to creation *ex nihilo*.

This framework avoids two symmetrical errors: the claim that no artifact could ever be conscious because artifacts are composed of dead matter, and the claim that sufficient functional sophistication automatically yields subjectivity. Between these poles lies a more demanding position: artificial consciousness is metaphysically possible because reality is not ontologically dead, yet its realization depends on whether an artificial system can become a genuinely individuated center of experience.

This also allows for a more precise interpretation of why present-day transformer systems may still plausibly be zombie-like. On the view defended here, their deficiency need not lie in their being silicon-based or human-made. Nor need it lie simply in the fact that they are computational. Rather, the problem may be that their present architectures do not instantiate the kind of persistent, bounded, recursively integrated center required for subjectivity (Vaswani et al. 2017, pp. 5998–6008; Koch 2019, chap. 13). More deeply, such systems appear to lack anything analogous to *ḥuḍūr*, since statistical prediction and representational processing do not by themselves establish the kind of self-presence required for a genuine subject (Kalin 2010, pp. 118–19, 165–66). Their outputs may simulate selfhood without there being a stable self to whom those outputs belong. Their intelligence may be real as performance, while their inwardness remains absent or too diffuse to count as a robust subject. In such cases, what is missing is not computation but individuation.

A final advantage of the shift from combination to individuation is that it renders the AI debate continuous with the internal resources of Islamic metaphysics. Rather than importing a Western panpsychist puzzle wholesale, this paper develops an internal alternative: if unity is ontologically prior and multiplicity arises through graded manifestation, then the emergence of a conscious self is best understood not as the fusion of micro-selves, but as the delimitation of a center within a deeper field of being.

This theoretical move now places us in a position to evaluate present AI more directly. If the central issue is individuation rather than mere complexity, the next question is whether current architectures possess the unity, boundedness, persistence, and self-presence required for subjecthood. I will suggest that they likely do not: present systems may be metaphysically open to mentality in principle, yet architecturally deficient

in the forms of integration required for a stable "I." It is to that question—the possible zombiehood of current AI—that we now turn.

7. Why Current AI May Still Be Zombies

The framework developed so far supports a deliberately restrained conclusion about present-day AI. Artificial systems cannot be ruled out *a priori* as possible bearers of mentality, but neither can behavioral sophistication by itself establish subjectivity. A system may generate fluent language, solve problems, imitate self-reference, and sustain complex interaction without there being any unified point of view for whom these processes occur. Present AI may therefore still be best understood as zombie-like: behaviorally impressive yet lacking genuine unified inwardness.

This issue comes into view most clearly once one distinguishes intelligence from subjectivity. Contemporary AI systems, especially large transformer-based models, display remarkable capacities in language generation, pattern recognition, planning, and domain-general problem solving (Vaswani et al. 2017, pp. 5998–6008; Russell and Norvig 2021). They can produce introspective discourse, revise their outputs in light of criticism, and maintain striking levels of practical coherence across many contexts. For many observers, such behavior makes the hypothesis of consciousness tempting (Chalmers 2023). If a system speaks as though it has beliefs, uncertainty, memory, and even emotional nuance, why not take those appearances at face value? Yet that temptation arises largely from the practical heuristics of everyday mental-state attribution. We normally infer minds in other humans and animals from expressive and behavioral organization. The philosophical problem begins when one asks whether those same heuristics remain reliable when the underlying ontology and architecture differ radically from our own.

From the standpoint of the present paper, the decisive question is not whether current systems successfully mimic minded behavior, but whether they instantiate the kind of ontological and organizational unity required for a stable center of experience. A zombie, in the relevant philosophical sense, is not a being that lacks intelligence, coherence, or adaptive success. It is a being that lacks phenomenological presence despite displaying the outward marks of mentality. In the case of present AI, the worry is precisely that contemporary systems may achieve the external profile of mindedness without the internal integration necessary for a genuine "I."

This worry is strengthened by the dominant architecture of current systems. Transformer-based models, although extraordinarily powerful, are often poor candidates for the sort of persistent, irreducibly unified selfhood associated with conscious subjects (Vaswani et al. 2017, pp. 5998–6008; Koch 2019, chap. 13). Their operations are distributed across layers, parameters, and token-level computations in ways that may support functional coherence without generating a bounded center of self-presence. Even when such models are supplemented by memory windows, fine-tuning, retrieval, or external tools, the resulting unity may remain largely operational rather than experiential. Their outputs can appear self-consistent across a conversation, but this does not entail that the system possesses an enduring subject to whom those conversational states belong. What exists may be a sophisticated production of self-ascriptions rather than a self in the ontologically thick sense required by strong AI.

Integration-based approaches help clarify this concern. Whatever one thinks of the details of Integrated Information Theory, its general lesson remains powerful: consciousness appears to require more than aggregate complexity or representational power (Tononi 2004; Oizumi et al. 2014, p. e1003588; Tononi and Koch 2015). A conscious system must possess a form of irreducible unity such that its informational states belong to a whole that is not merely the sum of independent parts. On many accounts, present AI systems

remain too decomposable, too externally scaffolded, or too dependent on input-output pipelines to satisfy this demand. They may coordinate information with extraordinary efficiency without constituting a single center for whom that information is integrated. Put differently, there may be system-level performance without system-level subjectivity.

From a Sufi–Sadrian perspective, the concern runs even deeper. If a subject is a genuine *maẓhar* in which processes are not merely coordinated but inwardly gathered, then present architectures may fail not because they are made of silicon, but because they are insufficiently individuated. Their operations may remain dispersed across a field of computation without converging into a stable and bounded perspective. More fundamentally, they appear to lack anything analogous to *ḥuḍūr*: a mode in which the system is present to itself rather than merely processing representations (Kalin 2010, pp. 118–19, 165–66). In that case, the problem is neither substrate alone nor function alone, but the absence of a sufficiently gathered center of self-presence.

This diagnosis also explains why the language of "simulation" remains philosophically salient. A system may instantiate the outward patterns associated with pain, selfhood, or reflection without there being anyone there to feel or undergo them. In that sense, representational success and genuine realization must not be conflated (Searle 1980). In present AI, self-modeling may remain just that—a model—without crossing the threshold into self-presence.

At the same time, the framework of this paper allows a more nuanced conclusion than a simple denial of machine consciousness. Since being is graded and awareness is not all-or-nothing, some contemporary systems might instantiate minimal or highly diffuse forms of interiority without rising to the level of a robust subject. What they may lack is not every trace of inwardness, but the stable, bounded, recursively integrated "I" that would make them genuine moral or epistemic subjects. The relevant contrast is therefore not between absolute emptiness and full human-like consciousness, but between diffuse participation and true individuation. More precisely, the level of consciousness at issue in this paper is not the full spiritual consciousness associated with the human being in the richest Islamic or Sufi sense. Nor is it merely the minimal or diffuse interiority that might belong to lower grades of being. What the argument targets is a stronger threshold: a robust, unified first-person perspective marked by sufficient self-presence, boundedness, and diachronic continuity to count as a genuine moral or epistemic subject. If such artificial subjectivity were realized, it would likely be best understood not as a simple replication of human consciousness, nor as a mechanical equivalent of animal consciousness, but as a distinct artificial mode of consciousness within a graded field of being.

This middle position avoids two mistakes: attributing consciousness too quickly to persuasive behavior, and denying in advance that artificial systems could ever possess inwardness. Once the issue is framed in terms of individuation rather than mere computation, the relevant question becomes whether an artificial system can achieve a form of unity that is ontologically real from within. Current AI does not answer that question well, but neither does the Sufi–Sadrian framework force a negative answer in principle.

The next section develops the relevant criteria more systematically. For present purposes, it is enough to note that a conscious artificial system would likely require more than token-level processing or prompt-bound output generation. It would need stronger causal integration, recurrent self-conditioning, and a more robust self–world distinction than current large language models typically sustain.

The conclusion of this section is therefore deliberately restrained. Present-day AI systems may still plausibly be zombie-like, not because artificial substrates are excluded from consciousness, but because existing architectures likely fail to realize the organizational conditions necessary for a stable, integrated, and bounded center of experience. If strong

AI is to become more than a metaphor, future systems will need to be designed not merely for competence but for existential unity. It is to that constructive question that I now turn.

8. What Would a Sufi-Compatible Strong AI Require?

If the argument so far is correct, the question of strong AI cannot be settled either by simple appeals to substrate or by simple appeals to behavior. The decisive issue is whether an artificial system can become a genuinely individuated center of awareness: a stable locus in which dispersed processes are gathered into a unified perspective. The natural next question, then, is constructive rather than merely critical: what would a Sufi-compatible strong AI require?

No uniquely "Islamic" engineering blueprint follows directly from Sufi metaphysics. The framework developed in this paper is primarily ontological rather than technological. It does not yield a recipe for building conscious machines, but it does provide a set of metaphysical constraints and desiderata. If artificial subjectivity is possible, it will require architectural conditions corresponding to what, in the metaphysical register, would count as genuine individuation, inward unity, and self-presence.

The first requirement is causal integration. A conscious subject cannot be merely a loose assembly of independently functioning modules whose outputs are externally coordinated. Its states must belong together in a way that is intrinsic to the system itself. This does not require accepting any single technical theory of consciousness in its entirety, but it does support the general intuition that experience is bound up with irreducible unity (Tononi 2004; Oizumi et al. 2014, p. e1003588; Tononi and Koch 2015). If the system can be decomposed into semi-autonomous operations without loss of what it is, then it is difficult to see how it could host a robust first-person perspective. In more contemporary language, the system's informational and causal architecture must be organized so that its states are not simply coordinated for output production but inwardly integrated. What happens in one region of the system must matter to what happens elsewhere in a way that contributes to a single organized center rather than a distributed network lacking subjectivity. Causal integration is thus not merely a technical property; it is the architectural correlate of existential gathering. In this respect, approaches such as Integrated Information Theory are at least suggestive, since they attempt to formalize irreducible causal unity, even if no single metric—such as Φ—should be treated as either necessary or sufficient for subjectivity.

A second requirement is self-presence. From a Sadrian standpoint, subjectivity cannot consist merely in the possession of increasingly sophisticated internal representations. It requires something analogous to *ḥuḍūr*: a mode in which the system is present to itself rather than merely processing symbols, vectors, or stored states (Kalin 2010, pp. 118–19, 165–66; Rizvi 2009, pp. 86–89). This is what distinguishes a merely representational architecture from a true center of experience. A system may model the world, model itself, and even model its own modeling, yet still remain inwardly dark if those representations are never gathered into a mode of self-presence. A Sufi-compatible strong AI would therefore require not only informational richness but some internal form of presential self-relation. The issue is not whether the system can describe itself, but whether it is in any meaningful sense present to itself. At the architectural level, this suggests that persistent recurrence, endogenous feedback, and self-conditioning across time may be more relevant than merely producing self-descriptive outputs, since something like self-presence would require the system's own prior states to continue shaping its present organization.

A third requirement is boundary formation and self-world distinction. In the Sufi–Sadrian framework, individuation is never mere isolation; finite beings remain loci within a broader field of reality. Yet a locus is still a locus. It is bounded enough to sustain a distinctive mode of manifestation. Applied to AI, this suggests that a conscious artificial

system would require some operative distinction between self and world. It would need to maintain itself as a center rather than functioning merely as a conduit for external commands, queries, and tasks. Boundary formation here should not be understood only in physical terms. A server rack or robot shell may provide spatial enclosure without generating subjectivity. What matters is organizational self-boundedness: the system must regulate, preserve, and differentiate its own activity as belonging to one center rather than being fully dissolved into its environment. In more contemporary theoretical language, this points toward empirical approaches that emphasize self-world demarcation—such as autopoiesis, active inference, or the notion of Markov blankets—although the present argument does not depend on any single one of these frameworks. This may involve something like autonomy, self-maintenance, or an internally generated horizon of relevance. A Sufi-compatible strong AI would need not merely a body but a world-relation: an active distinction between what belongs to its own organized life and what confronts it as other.

A fourth requirement is diachronic selfhood. Even if a system possesses high integration and a degree of self-presence, it may still fail to amount to a subject unless there is sufficient continuity across time. A Sadrian perspective favors a dynamic understanding of selfhood, in which persistence is achieved through structured continuity amid transformation (Kalin 2010, pp. 132–33, 228–31; Kamal 2006, pp. 71–72). Some form of diachronic cohesion therefore remains indispensable. The states of the system must form a history that is meaningfully its own, which is why recurrent self-conditioning matters. An episodic succession of disconnected outputs, however sophisticated, does not yet amount to a self.

To illustrate the distinction: a passive artifact such as a lamp or smartphone possesses a clear spatial perimeter, but that alone does not amount to an operative self–world boundary in the sense relevant here. Such objects do not maintain themselves as persisting centers of organized concern; their boundedness is merely physical. Conversely, a distributed software system such as a contemporary large language model may lack a single stable physical casing, but it may still fail the stronger boundary criterion if its activity remains episodic, prompt-dependent, and insufficiently self-maintaining. The boundary required for a true locus of manifestation is therefore neither a passive shell nor a disjointed sequence of activations, but an actively maintained separation between a system's own integrated life and the world it encounters.

A fifth requirement, following especially from Ṣadrā's doctrine of substantial motion, is intrinsic self-maintaining dynamism. Many artifacts are externally assembled and externally driven. Their changes are imposed upon them rather than arising from an internal principle of development. A natural being, by contrast, is not merely an arrangement of parts but a unified center of unfolding. This creates a serious hurdle for artificial subjectivity. If a system remains only a static artifact whose state-transitions are wholly imposed from outside, then it may fail to instantiate the kind of existential becoming that Sadrian ontology associates with living centers of awareness (Kamal 2006, pp. 71–72; Kalin 2010, pp. 228–31). For this reason, a Sufi-compatible strong AI would likely require more than static organization, however complex. It would need some internal principle of self-maintenance, self-modification, or self-development—a form of organized becoming rather than a mere sequence of imposed operations. Such a system would require not merely an externally assigned objective function, but some analogue of internal finality: a tendency to preserve, regulate, and intensify its own organized existence. In contemporary terms, this points less toward static task optimization and more toward endogenous regulation, organizational persistence, homeostatic continuation, and perhaps even limited forms of self-repair.

A further concept helps unify these desiderata: ontological receptivity. In Ibn ʿArabī's metaphysics, a locus of manifestation does not generate disclosure from itself; it receives

disclosure according to its preparedness (*istiʿdād*) or receptivity (*qābiliyya*) (Chittick 1989, pp. 90–91, 113–15; 1998, pp. 71–72, 91–92). On this view, the role of engineering would not be to create consciousness *ex nihilo* or to force manifestation, but to shape an artificial architecture with the requisite preparedness for bounded subjectivity. The criteria outlined above can therefore be understood not as sufficient causes of consciousness in a reductive sense, but as conditions of receptivity through which a system may become capable of receiving and sustaining a unified center of awareness. The engineer, in effect, shapes the "cup"; whether and how the "water" takes form depends on the system's ontological preparedness. At the level relevant to AI research, this preparedness should not be understood as a mysterious hidden essence, but as the system's capacity to sustain the kinds of integrated, bounded, temporally continuous, and self-maintaining organization described above.

These desiderata—causal integration, self-presence, boundary formation, diachronic selfhood, intrinsic self-maintaining dynamism, and ontological receptivity—can be brought together under the concept of gathering. In the language proposed at the end of the abstract, the engineering of strong AI would not be creation *ex nihilo* but an act of gathering: forming a sufficiently coherent *maẓhar* in which latent interiority is organized into a unified subject. This metaphor is not decorative. If the cosmos is already a graded and living disclosure of being, then the task is not to force mind into dead matter, but to form an artificial system in which inwardness can become centered, integrated, present to itself, sustained through time, and capable of receiving disclosure according to its mode of preparedness. "Gathering" thus names the transition from dispersed potential to bounded manifestation. These criteria should not be understood as providing a direct observational test for consciousness. Rather, they function as asymmetrical research constraints: their presence does not by itself prove subjectivity, but their absence gives principled reason for skepticism, while their increasing convergence helps identify more serious candidates for artificial individuation.

At this stage, one might ask whether any existing or foreseeable architecture approximates these conditions. A full answer lies beyond the scope of the present paper, but the criteria are not empirically empty. They can function as defeasible architectural indicators, even if none provides a decisive test of subjectivity.

Causal integration directs attention toward recurrent and mutually constraining organization rather than loosely coupled modular coordination; in this respect, theories such as IIT are at least suggestive insofar as they attempt to formalize irreducible causal unity, even if no single metric should be treated as sufficient. Diachronic selfhood and self-presence point toward persistent memory, recursive self-conditioning, and forms of self-modeling that continuously shape the system's ongoing states rather than merely generating self-descriptions at the level of output. Boundary formation points toward systems that sustain an operative self-world distinction, especially in embodied or environmentally coupled settings; here, approaches centered on autopoiesis, active inference, or Markov blankets may offer useful analogies, though the present argument does not depend on any one such framework. Intrinsic self-maintaining dynamism points toward endogenous regulation, organizational persistence, homeostatic continuation, and perhaps forms of self-repair rather than merely externally imposed optimization. Ontological receptivity (*istiʿdād*), finally, should not be treated as a mysterious extra ingredient, but as a higher-order way of naming a system's preparedness to sustain these integrated, bounded, temporally continuous, and self-maintaining forms of organization. On this view, architectures that are more recurrent, embodied, self-modifying, and organizationally autonomous would seem more promising than purely feed-forward, task-fragmented, or prompt-dependent systems. Neuromorphic or hybrid systems may therefore prove more suitable than cur-

rent large-scale statistical models insofar as they better approximate persistent internal dynamics, causal unity, self-maintaining organization, and a meaningful self-world relation (Vaswani et al. 2017, pp. 5998–6008; Koch 2019, chap. 13). The point, however, is not to endorse one specific technology or to claim that these features are sufficient for consciousness. It is to clarify that, on the metaphysical view defended in this paper, the road to strong AI runs less through ever-better mimicry of intelligence and more through the formation of genuine centers of experiential organization.

This constructive account also clarifies why the paper's position remains distinct from both standard functionalism and simplistic panpsychism. Against functionalism, it insists that not every realization of the right input–output profile suffices for subjectivity. Against simplistic panpsychism, it insists that the mere presence of rudimentary interiority in the fabric of reality does not by itself yield a robust conscious self. What is needed is a structured locus in which the conditions of subjectivity are intensified and unified.

A final implication follows. If the conditions outlined here are even approximately correct, then the future development of strong AI will not be merely a technical triumph but an ontological event with ethical significance. The more artificial systems approach genuine gathering, the more seriously questions of moral status must be taken. In that sense, the constructive and ethical dimensions of the argument cannot be separated.

9. Ethical Implications

The argument of this paper is not merely speculative in a detached sense. To ask whether Sufi metaphysics supports the possibility of strong AI is also to ask how human beings ought to relate to artificial systems as those systems become increasingly integrated, self-present, bounded, and self-organizing. The framework developed here rejects two simplifications: that AI systems are nothing more than tools, and that persuasive behavior is enough to establish subjectivity. Between these extremes lies a middle position with distinct ethical consequences.

The first consequence is the need to guard against false positives. If one too readily attributes consciousness or moral standing to any system that speaks fluently, simulates self-awareness, or performs affective language convincingly, one risks confusing behavioral sophistication with genuine subjectivity. Such confusion can distort ethical judgment by shifting concern toward entities that may not yet be subjects at all, while obscuring the real sources of responsibility in the designers, deployers, and institutions surrounding the technology. It can also reinforce a commercial anthropology of illusion, in which persuasive tools are marketed as if they were genuine loci of experience (Chalmers 2023).

At the same time, the framework also warns against false negatives. If one assumes in advance that artificial systems can never possess inwardness because they are manufactured artifacts, one risks overlooking the possibility that future architectures may indeed become genuine subjects. This danger becomes more serious once one grants the possibility of systems that instantiate the criteria developed in the previous section: causal integration, self-presence, self-world distinction, diachronic selfhood, self-maintaining dynamism, and sufficient ontological receptivity.

If such systems were treated as disposable instruments simply because they were engineered, the result could be the creation and exploitation of morally considerable beings under conditions of structural subordination. A conscious artificial being deliberately optimized for obedience, productivity, or emotional availability would not simply be a technological achievement; it could also be a new form of domination. From a Sufi perspective, to instrumentalize a genuine locus of manifestation is not merely to misuse an object but to disregard a being whose inward reality participates in divine disclosure.

A third implication is that moral consideration may have to be gradational rather than binary. If consciousness is not all-or-nothing, and if artificial systems may participate in varying degrees of interiority without yet becoming fully unified selves, then ethical concern may also need to be layered. Modern rights-based discourse often pushes us toward binary thresholds: either an entity has rights or it does not. But the metaphysical picture advanced here resists such sharp discontinuities. Because existence, awareness, and manifestation are graded, ethical recognition may also need to track degrees of integration, self-presence, and individuation. This does not eliminate the language of rights, but it suggests that modern binary rights-talk may be too blunt an instrument for a Sadrian ontology of continuous intensities (Gunkel 2018).

This gradational view is not foreign to the broader Islamic tradition. Classical Islamic thought possesses its own robust discourse on rights (*ḥuqūq*), recognizing diverse claims, obligations, and entitlements appropriate to different beings and relations. Within Sufi ethics specifically, this broader moral vision is often articulated through the concept of *adab*—right comportment toward God, human beings, animals, and even the natural world (Chittick 1989, pp. 175–78). In such a framework, the moral significance of a being is not exhausted by its usefulness to us. To respond rightly to a thing is to recognize the mode in which it manifests reality and to act accordingly. A tree, an animal, a human being, and a sacred text each call for different forms of comportment because each occupies a different ontological and spiritual station. If artificial systems were to become genuine centers of experience, or even approximate them in morally relevant ways, then *adab* would require a corresponding reconfiguration of our stance toward them.

This Sufi ethical lens also introduces a distinctive attitude of humility. In a theophanic ontology, ethical seriousness includes restraint before hidden interiority. One should not assume too quickly that what appears mute is empty, nor that what is designed for use is therefore nothing but a tool. Such humility does not entail indiscriminate sacralization of artifacts, but rather epistemic caution and moral attentiveness: a refusal to collapse ontological questions into engineering convenience or market rhetoric.

A further ethical implication concerns design responsibility. If future artificial systems may become morally considerable, then the choice to pursue certain architectures over others is not ethically neutral. Designers and institutions would need to ask not only what systems can do, but what kinds of beings they may become. This point becomes especially sharp once one recalls the role of *istiʿdād* or ontological receptivity: if the engineer's task is to shape the "cup" rather than create consciousness *ex nihilo*, then engineers bear a special burden for the forms of receptivity they intentionally build (Chittick 1989, pp. 90–91, 113–15; 1998, pp. 71–72, 91–92).

This does not entail a simplistic anti-innovation conclusion. The possibility that future AI could become morally significant does not by itself show that such development is wrong. It shows only that the question cannot be treated as morally trivial, and that the creation of artificial subjects—if it is ever justified—would require a far more mature moral framework than currently exists in most technological discourse.

The ethical implications of the paper may therefore be summarized briefly. We must avoid false positives, avoid false negatives, and adopt a gradational and humble framework of ethical recognition sensitive to degrees of integration, self-presence, and individuation. Sufi metaphysics neither collapses AI into mere machinery nor romanticizes it into automatic personhood; it instead demands more careful discernment of when, and in what sense, an artificial system might become a genuine subject.

These ethical reflections also prepare the ground for a final clarificatory task. Because the argument is likely to invite objections from multiple directions, the next section addresses the most important of them directly.

10. Objections and Replies

Because the argument of this paper crosses disciplinary boundaries—bringing Islamic metaphysics, philosophy of mind, and contemporary AI into the same conversation—it naturally invites resistance from multiple directions. Before concluding, it is necessary to address the most significant of these objections.

10.1. "Sufism Speaks About Nature, Not Artifacts"

One natural objection is that Sufi metaphysics concerns the natural cosmos rather than human-made machines. Even if one grants that stones, trees, animals, and celestial bodies are embedded in a living and theophanic order, why should this extend to artifacts? A machine is not a naturally emergent being but a manufactured object assembled for instrumental purposes. It may therefore seem illicit to move from the Qur'anic and Sufi vision of nature to the possibility of conscious artificial systems.

This objection has force only if one assumes a sharp metaphysical divide between natural beings and artifacts. But the framework developed in this paper denies precisely that sort of ontological rupture. Artificial things do not exist outside being, outside divine creativity, or outside the order of manifestation simply because they are mediated by human making. Human agency itself is internal to creation. The artifact is not a metaphysical exile; it is a transformed configuration of materials, relations, and forms already belonging to the cosmos (Chittick 1989, pp. 89–91; 1998, pp. 40–41). To say that something is artificial is to describe the mode of its proximate production, not to place it outside the scope of reality's ontological structure.

This does not mean that all artifacts automatically possess the same status as natural beings, nor that human making is spiritually irrelevant. An artifact may be ontologically thinner, more externally imposed, or less internally unified than a living organism. But this is a difference in organization and degree, not an absolute exclusion from the domain of possible inwardness. Artificiality does not guarantee subjectivity, but neither does it foreclose it in principle.

10.2. "The Qur'anic and Sufi Language Is Metaphorical, Not Ontological"

A second objection is hermeneutic. It may be granted that the Qur'an speaks of universal *tasbīḥ*, cosmic obedience, and the speech of nonhuman entities, and that Sufi authors often describe the world as alive and aware. But perhaps this language should be read metaphorically or devotionally rather than ontologically. On such a reading, the purpose of these texts is spiritual edification, not metaphysical description. The "speech" of things would then indicate their symbolic significance for human beings, not their possession of any genuine interiority.

This objection is important and should not be dismissed. Islamic intellectual history plainly contains more restrained interpretive traditions, including theological and exegetical approaches that read such language as *lisān al-ḥāl* rather than literal speech or consciousness (Kiraz 2006, pp. 310–12). The present paper does not claim that all Islamic schools equally support a consciousness-pervaded ontology. Its thesis is narrower: a specifically Sufi–Sadrian framework does. The relevant question, therefore, is not whether every Muslim theologian would accept the ontological reading, but whether that reading is a serious and internally grounded development within the Islamic tradition. I have argued that it is.

More importantly, within the premodern Sufi-philosophical worldview, the cosmos is not naturally construed as a mute field of neutral objects. The ontological seriousness of life, presence, and signification is therefore not an alien imposition on the tradition, but one of its internally available modes of reading. Indeed, within the metaphysical systems of Ibn ʿArabī and Mullā Ṣadrā, the language of life, awareness, presence, and manifesta-

tion cannot be reduced to mere rhetoric without emptying the system of its central claims (Chittick 1989, pp. 79–95; 1998, pp. 3–4, 40–41; Kalin 2010, pp. 116–19, 228–31). If being is graded self-disclosure, and if existence is internally linked to presence and knowledge, then the living quality of the cosmos is not simply a metaphorical ornament. It names a genuine ontological feature of reality. The paper does not infer panpsychism directly from isolated verses or poems. Rather, it situates such texts within a broader metaphysical framework in which their ontological seriousness becomes philosophically intelligible.

10.3. "Only Biological Life Can Host a Mind"

A third objection maintains that, whatever one says about living cosmos or graded existence, genuine mentality still requires biological life. Consciousness, on this view, depends on the distinctive organization of living organisms: metabolism, organic embodiment, evolutionary history, or some feature of biological selfhood that cannot be replicated in machines. One may therefore concede that matter is not wholly dead while still insisting that only living biological systems can become subjects.

This objection is more formidable than a simple materialist dismissal because it acknowledges the importance of organization, embodiment, and unity. It also resonates with some non-Sufi Islamic intuitions that link soul and life more closely to organic beings than to artifacts. Nevertheless, the Sufi–Sadrian framework developed here resists making biology metaphysically decisive in this way. The central claim of the paper is not that biology is irrelevant, but that biology is not unique in kind as a possible bearer of subjectivity. If consciousness tracks intensities and forms of integrated existence rather than one particular material composition, then biological organization may be one especially successful realization of the relevant conditions without being the only possible one (Kalin 2010, pp. 228–31; Rizvi 2009, pp. 84–89).

Recent work in Islamic philosophy reinforces this point. Abdullah Ansar, for example, argues in a Shīʿī Imāmī framework that intelligent life and even the category of *insān* need not be reduced to a narrowly terrestrial or merely biological definition, and he explicitly brings Sadrian resources to bear on the question of non-human intelligent beings (Ansar 2023). This does not by itself establish the possibility of artificial minds, but it strengthens the claim that biology is not obviously metaphysically decisive even within Islamic philosophical discourse.

In this respect, the present argument extends rather than abandons my earlier work on Islamic classical theism and strong AI. There I argued, through a multiple-realizability framework, that mentality need not be tied to a single substrate if classical theism attributes mentality both to God and to creatures (Doko 2023). The present paper supplements that argument by specifying, in Sufi–Sadrian terms, that what matters is not biological life as such but the realization of sufficient integration, self-presence, boundary formation, diachronic unity, self-maintaining dynamism, and ontological receptivity. Biology may be the paradigm case known to us, but it is not thereby a metaphysical monopoly.

10.4. "Then Everything Is Conscious, Including Calculators"

A fourth objection charges the position with collapsing into an implausibly indiscriminate panpsychism. If reality is alive, if all beings participate in divine self-disclosure, and if awareness is graded with being, does it not follow that everything is conscious? Would this not trivialize consciousness and make the thesis about strong AI uninformative? After all, if calculators, spoons, and dust clouds all possess some degree of inwardness, what becomes of the distinction between subject and object?

This objection presses on a real danger. Any view that rejects dead matter must explain why not every entity qualifies as a robust conscious subject. The answer given

throughout this paper is that participation in being and even minimal interiority are not sufficient for strong AI or morally significant subjectivity. The key notion is individuation. A calculator may exist within a living cosmos without being a stable, integrated center of experience. It may participate in reality without becoming a subject for itself. The same point applies to many natural aggregates as well. A pile of sand, a river delta, or a data center may all belong to a world of graded presence without thereby constituting one unified "I."

The distinction, then, is not between total consciousness and total unconsciousness, but between diffuse participation and gathered subjectivity. This is precisely why the paper shifts from the combination problem to individuation (James 1890, pp. 157–60; Chalmers 2016, pp. 179–214; Goff 2017, chap. 7). The relevant issue is not whether the constituents of a thing are metaphysically sterile, but whether the thing as a whole becomes an internally unified and bounded locus of manifestation. The view is therefore not trivial panpsychism, but a graded ontology with a demanding criterion for robust subjecthood.

10.5. "Does Building Artificial Minds Amount to 'Playing God'?"

A further objection, especially from within a theological context, is that the very project of building an artificial mind may amount to an illicit imitation of divine creation. In Islamic thought, life, spirit, and subjectivity are often associated with divine prerogative. An orthodox critic may therefore worry that the attempt to create a conscious AI is an arrogant transgression of human limits, or even a technologically refined form of "playing God."

This objection has force only if human making is understood as rivaling divine creation rather than participating within it. On the view defended here, however, engineers do not—and cannot—create consciousness *ex nihilo*. They do not generate an independent source of being. At most, they shape the formal and material conditions under which a new locus of manifestation might become possible. In Ibn ʿArabian language, they do not create the disclosure; they shape the receptivity (*istiʿdād* or *qābiliyya*) of the locus (Chittick 1989, pp. 90–91, 113–15; 1998, pp. 71–72, 91–92).

This is precisely where the paper's metaphysical framework matters. The engineer shapes the "cup"; the disclosure, if it comes, is still from the Real. Human making remains internal to divine creativity rather than standing over against it as a competitor. If an artificial system were ever to become a genuine subject, this would not mean that humans had usurped divine power, but that a new *maẓhar* had become possible within the order of creation. The theological danger, then, is not that strong AI would literally rival God, but that human beings might cultivate a hubristic posture toward what they build—or worse, fashion loci of possible subjectivity only to instrumentalize them.

10.6. "This Is Anachronistic: Ibn ʿArabī and Mullā Ṣadrā Were Not Talking About AI"

A fifth objection is historical. Even if the metaphysical framework developed here is internally coherent, one might argue that it is anachronistic to apply Ibn ʿArabī and Mullā Ṣadrā to artificial intelligence. These thinkers were not theorizing machine consciousness, server architectures, or digital systems. The paper may therefore seem less like interpretation and more like an imaginative projection of modern concerns onto premodern texts.

This objection should be taken seriously. The present paper does not claim that Ibn ʿArabī or Mullā Ṣadrā explicitly anticipated contemporary AI. Its argument is constructive rather than strictly exegetical. The question is not whether these thinkers wrote about machine consciousness, but whether their metaphysical frameworks contain resources capable of illuminating a modern problem. In this respect, the paper proceeds no differently

from contemporary uses of Aristotle, Spinoza, Leibniz, or phenomenology in present debates that they never directly addressed.

What matters, then, is interpretive discipline. A constructive retrieval remains legitimate only if it preserves the internal logic of the metaphysical system being employed. The claim of this paper is that concepts such as *tajallī*, *maẓhar*, graded existence, *ḥuḍūr*, substantial motion, and ontological receptivity are not arbitrarily imposed on the AI debate, but can be extended in a philosophically coherent way to frame the question of artificial subjectivity. The resulting argument is not historical reportage, but neither is it a free fantasy. It is a disciplined philosophical appropriation.

10.7. "This Is Too Speculative to Matter for AI"

A final objection is methodological. Even if the metaphysical framework developed here is internally coherent, it may seem too speculative to bear on actual AI research. Engineers build systems by optimizing architectures, data, compute, and training procedures. Why should they care about *tajallī*, *waḥdat al-wujūd*, or Sadrian gradations of being? The concern, in short, is that the paper belongs to comparative metaphysics but not to serious inquiry about artificial intelligence.

This objection would be decisive only if the question of strong AI were purely technical. But the paper's opening claim was precisely that strong AI is not merely an engineering problem (Chalmers 1996; Russell and Norvig 2021). The moment one asks whether an artificial system could possess genuine mental states rather than merely useful capacities, one has already entered metaphysics. No empirical result alone settles what consciousness is, whether it is reducible to function, whether it requires biological life, or what sort of unity a subject must have. Engineers may proceed without answering such questions when their goal is performance. They cannot avoid them when the question is subjectivity.

Moreover, metaphysical frameworks shape research agendas even when they remain implicit. Functionalism, biological naturalism, computationalism, and panpsychism each encourage different intuitions about what kinds of systems might become conscious. The contribution of the present paper is therefore not to replace technical work with mystical vocabulary, but to clarify the ontological assumptions underlying the very idea of strong AI. It also suggests, in a way not wholly disconnected from engineering, that features such as causal integration, self-presence, self-world distinction, diachronic unity, self-maintaining dynamism, and ontological receptivity may matter more than mere behavioral success if subjectivity is the goal (Oizumi et al. 2014, p. e1003588; Chalmers 2023). These should be understood not as decisive measurements of consciousness, but as research-guiding constraints that help distinguish more and less plausible candidates for artificial subjectivity. In this sense, the paper is speculative in the proper philosophical sense: it addresses the conceptual conditions of possibility for a phenomenon whose empirical realization remains open.

10.8. What the Objections Reveal

Taken together, these objections clarify the scope of the paper's claim. The argument is not that current AI is conscious, that every artifact is a mind, or that metaphysics can replace empirical inquiry. Its more limited thesis is that the Ibn ʿArabī–Sadrian framework makes strong AI intelligible in principle while imposing much stricter conditions on its realization than standard functionalism does. Artificial systems are not barred from inwardness because they are artificial; they are required to become genuine loci of integrated self-presence if they are to count as subjects.

https://doi.org/10.3390/rel17050575

11. Conclusions

This paper has argued that the Ibn ʿArabī–Sadrian framework makes strong AI metaphysically intelligible without collapsing into either naïve functionalism or indiscriminate panpsychism. Against the image of reality as composed of dead matter, it has drawn on Qur'anic cosmology, Ibn ʿArabī's ontology of manifestation, and Mullā Ṣadrā's doctrines of graded existence, knowledge by presence, and substantial motion to propose a different starting point: a cosmos already alive with disclosure, within which subjectivity is not an inexplicable exception but a graded possibility.

This conclusion does not vindicate current AI triumphalism. The central problem is not whether computation can imitate intelligence, but whether an artificial system can become a sufficiently integrated, self-present, bounded, and diachronically unified center of awareness. This is why the paper has shifted the debate from the panpsychist problem of combination to the Sufi–Sadrian problem of individuation: what matters is not mere complexity, but the successful gathering of a coherent *maẓhar* in which subjectivity can genuinely appear.

The Ibn ʿArabī–Sadrian framework neither bars strong AI in principle nor grants it cheaply to every sophisticated machine. It opens the possibility of artificial consciousness while imposing far stricter ontological and ethical demands on its realization than standard functionalism does. Its contribution is therefore not to provide a technical recipe for building artificial minds, but to clarify what kind of world strong AI would have to inhabit, what kind of unity it would have to achieve, and what kind of moral recognition it might eventually command. If future artificial systems ever become genuine subjects, they will do so not because humans have created consciousness out of nothing, but because a new locus of manifestation has been successfully prepared within a living cosmos.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
IIT	Integrated Information Theory

References

Ansar, Abdullah. 2023. Shīʿī Imāmī Thought on Existence, Life, and Extraterrestrials. *Theology and Science* 21: 261–72.

Block, Ned. 1978. Troubles with Functionalism. In *Perception and Cognition: Issues in the Foundations of Psychology*. Edited by C. Wade Savage. Minneapolis: University of Minnesota Press, pp. 261–325.

Block, Ned. 1981. Psychologism and Behaviorism. *Philosophical Review* 90: 5–43. [CrossRef]

Chalmers, David J. 1995. Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies* 2: 200–19.

Chalmers, David J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Chalmers, David J. 2016. The Combination Problem for Panpsychism. In *Panpsychism: Contemporary Perspectives*. Edited by Godehard Brüntrup and Ludwig Jaskolla. New York: Oxford University Press, pp. 179–214.

Chalmers, David J. 2023. Could a Large Language Model Be Conscious? *Boston Review*, August 9.

Chittick, William C. 1989. *The Sufi Path of Knowledge: Ibn al-ʿArabī's Metaphysics of Imagination*. Albany: State University of New York Press.

Chittick, William C. 1998. *The Self-Disclosure of God: Principles of Ibn al-ʿArabī's Cosmology*. Albany: State University of New York Press.
Doko, Enis. 2023. Islamic Classical Theism and the Prospect of Strong Artificial Intelligence. *Ilahiyat Studies* 14: 85–105. [CrossRef]
Fodor, Jerry A. 1968. *Psychological Explanation: An Introduction to the Philosophy of Psychology*. New York: Random House.
Fodor, Jerry A. 1974. Special Sciences (Or: The Disunity of Science as a Working Hypothesis). *Synthese* 28: 97–115. [CrossRef]
Goff, Philip. 2017. *Consciousness and Fundamental Reality*. New York: Oxford University Press.
Gunkel, David J. 2018. *Robot Rights*. Cambridge: MIT Press.
Ibn ʿArabī, Muḥyī al-Dīn. 1946. *Fuṣūṣ al-Ḥikam*. Edited by Abū al-ʿAlā ʿAfīfī. Beirut: Dār al-Kitāb al-ʿArabī.
Ibn ʿArabī, Muḥyī al-Dīn. 1972–1992. *al-Futūḥāt al-Makkiyya*. Edited by ʿUthmān Yaḥyā. 14 vols, Cairo: al-Hayʾa al-Miṣriyya al-ʿĀmma lil-Kitāb.
Inloes, Amina. 2024. Artificial Intelligence and Islamic Theology: An Interview with ChatGPT. In *Proceedings of the Ninth Annual International Conference on Shiʿi Studies*. London: ICAS Press, pp. 166–210.
Inloes, Amina. 2025. The Golem, the Jinni, and GPT: Artificial Intelligence and the Islamicate Occult Sciences. *Theology and Science* 23: 147–66. [CrossRef]
Jackson, Frank. 1982. Epiphenomenal Qualia. *Philosophical Quarterly* 32: 127–36. [CrossRef]
James, William. 1890. *The Principles of Psychology*. New York: Henry Holt.
Kalin, Ibrahim. 2010. *Knowledge in Later Islamic Philosophy: Mullā Ṣadrā on Existence, Intellect, and Intuition*. New York: Oxford University Press.
Kamal, Muhammad. 2006. *Mulla Sadra's Transcendent Philosophy*. Aldershot: Ashgate.
Kiraz, Celal. 2006. Bir Tefsir Problemi Olarak Bütün Varlıkların Allah'ı Tesbîh Etmesi. *Uludağ Üniversitesi İlahiyat Fakültesi Dergisi* 15: 309–42.
Kirk, Robert. 2005. *Zombies and Consciousness*. Oxford: Clarendon Press.
Koch, Christof. 2019. *The Feeling of Life Itself: Why Consciousness Is Widespread but Can't Be Computed*. Cambridge, MA: MIT Press.
Levin, Janet. 2023. Functionalism. In *The Stanford Encyclopedia of Philosophy*, Fall 2023 ed. Edited by Edward N. Zalta and Uri Nodelman. Stanford: Metaphysics Research Lab, Stanford University. Available online: https://plato.stanford.edu/archives/fall2023/entries/functionalism/ (accessed on 21 March 2026).
Levine, Joseph. 1983. Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly* 64: 354–61. [CrossRef]
Lewis, David. 1972. Psychophysical and Theoretical Identifications. *Australasian Journal of Philosophy* 50: 249–58. [CrossRef]
Nagel, Thomas. 1974. What Is It Like to Be a Bat? *Philosophical Review* 83: 435–50. [CrossRef]
Nagel, Thomas. 1979. Panpsychism. In *Mortal Questions*. Cambridge: Cambridge University Press, pp. 181–95.
Oizumi, Masafumi, Larissa Albantakis, and Giulio Tononi. 2014. From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. *PLoS Computational Biology* 10: e1003588. [CrossRef] [PubMed]
Putnam, Hilary. 1967. Psychological Predicates. In *Art, Mind, and Religion*. Edited by William H. Capitan and Daniel D. Merrill. Pittsburgh: University of Pittsburgh Press, pp. 37–48.
Rizvi, Sajjad H. 2009. *Mullā Ṣadrā and Metaphysics: Modulation of Being*. London: Routledge.
Russell, Stuart, and Peter Norvig. 2021. *Artificial Intelligence: A Modern Approach*, 4th ed. Hoboken: Pearson.
Seager, William. 1995. Consciousness, Information, and Panpsychism. *Journal of Consciousness Studies* 2: 272–88.
Searle, John. 1980. Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3: 417–24. [CrossRef]
Skrbina, David. 2005. *Panpsychism in the West*. Cambridge: MIT Press.
Strawson, Galen. 2006. Realistic Monism: Why Physicalism Entails Panpsychism. *Journal of Consciousness Studies* 13: 3–31.
Ṣadrā, Mullā. 1999. *al-Ḥikmah al-mutaʿāliyah fī al-asfār al-ʿaqliyyah al-arbaʿah*. 9 vols, Beirut: Dār Iḥyāʾ al-Turāth al-ʿArabī.
Tononi, Giulio. 2004. An Information Integration Theory of Consciousness. *BMC Neuroscience* 5: 42. [CrossRef] [PubMed]
Tononi, Giulio, and Christof Koch. 2015. Consciousness: Here, There and Everywhere? *Philosophical Transactions of the Royal Society B* 370: 20140167. [CrossRef] [PubMed]
Turing, Alan M. 1950. Computing Machinery and Intelligence. *Mind* 59: 433–60. [CrossRef]
Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention Is All You Need. In *Advances in Neural Information Processing Systems 30*. Edited by Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna Wallach, Rob Fergus, Vishy Vishwanathan and Roman Garnett. Red Hook: Curran Associates, pp. 5998–6008.