

Article

# Connecting Consciousness to Physical Causality: Abhinavagupta's Phenomenology of Subjectivity and Tononi's Integrated Information Theory

Loriliai Biernacki

Religious Studies, University of Colorado, 292 UCB, Eaton Humanities 240, Boulder, CO 80301, USA; Loriliai.biernacki@colorado.edu; Tel.: +1-303-492-4730

Academic Editors: Glen A. Hayes and Sthaneshwar Timalsina

Received: 28 May 2016; Accepted: 23 June 2016; Published: 1 July 2016

**Abstract:** This article demonstrates remarkably similar methods for linking mind and body to address the “hard problem” in the work of 11th-century Indian philosopher Abhinavagupta with a currently prominent neuroscientific theory, Tononi's Integrated Information Theory 3.0. Both Abhinavagupta and Tononi and Christof Koch hinge their theories on the identity of phenomenal subjective experience with causality. Giulio Tononi's Integrated Information Theory is remarkable precisely in its method for dealing with the mind-body problem; namely, Tononi's mathematically oriented systems neurology proposes something we typically do not find in neuroscientific literature—that we start from a phenomenology of experience. Abhinavagupta's sophisticated and, for his milieu, novel way of linking subjectivity and objectivity in the concepts of knowledge (jñāna) and action (kriyā) also offers a way of understanding how subjectivity can be linked to causality. This particular configuration is mostly absent in Western Cartesian models for understanding consciousness and in Indian philosophical speculations on consciousness. However, this, in any case, is precisely the move that Tononi makes when he proposes that information is both “causal and intrinsic.” Abhinavagupta's similar linkage of subjectivity with causality can help us to think about Tononi's neuroscientific mathematical model.

**Keywords:** Abhinavagupta; cognitive science; Integrated Information Theory; Phi; consciousness

---

## 1. Connecting Consciousness to Physical Causality: Abhinavagupta's Phenomenology of Subjectivity and Tononi's Integrated Information Theory

This paper addresses methods for thinking about how to connect mind, as internal, phenomenological awareness, with physical causality in two culturally disparate models of consciousness—both of which, nevertheless, share a preoccupation with linking consciousness to causal efficacy through identifying the two. On the one hand, I draw from Tononi's Integrated Information Theory<sup>1</sup> and on the other, I use conceptions of consciousness from the medieval Indian thinker Abhinavagupta.

Tononi's Integrated Information Theory (IIT) has been recently gaining widespread support from cognitive scientists, notably prominent neuroscientist Christof Koch, who calls it “the best theory we have to date” [3]. In earlier work, Koch's position evinced an unqualified biological materialist reductionism [4]. His recent advocacy of Tononi's IIT marks a fundamental shift, with a departure

---

<sup>1</sup> While Tononi's *Phi: A Voyage from the Brain to the Soul* [1] offers a broad theoretical view of Integrated Information theory, much of my understanding of the precise and technical aspects of Tononi's theory of Phi derives from the most recent iteration of Information Integration Theory—IIT 3.0 in [2]. This source in PLOS is unapaginated; I pasted it into a word document and noted the page numbers from this and can supply them if helpful.

from materialist reductionism, to the extent that philosopher John Searle called it “anti-reductionism with a vengeance” ([5], p. 55).<sup>2</sup>

Apart from IIT’s increasing popularity as a viable theory for explaining consciousness, however, Tononi’s theory is remarkable precisely in its method of dealing with the mind-body problem. Tononi’s mathematically oriented systems neurology proposes something we typically do not find in neuroscientific literature—namely, that we start from a phenomenology of experience. From this phenomenology he proposes to derive mathematical principles for what kinds of systems display consciousness. The hopeful outcome on a merely practical level is for Tononi to develop a way for us to measure how much consciousness is present in, for instance, a comatose patient, in other animal species, even in computer artificial intelligence (AI) [2].

Key to his theory is the notion that information, when it is integrated, is itself the expression of consciousness. In this way, not explicitly addressing the “hard problem” of how mind and body relate, Tononi’s model in effect sidesteps this problem by its definition of information as both “causal and intrinsic” [2]. I will discuss at greater length below how this maps out, in Tononi’s theory and also by way of comparison with Abhinavagupta’s understandings of the components that make up consciousness. In any case, one of the most critiqued and controversial implications of Integrated Information Theory is the panpsychism it entails, which both Tononi and Koch unabashedly embrace. This shift in Koch’s position from a former adherence to a staunch materialist reductionism into panpsychism follows on the heels of some recent advocacy of panpsychism among philosophers of science, perhaps best exemplified by prominent philosopher of science Thomas Nagel’s recent controversial *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False* [6] even as we keep in mind that Chalmers, who coined the phrase “hard problem,” proposed a panpsychism as a means for explaining the mind-body connection as early as 1996 [7].

While Tononi proposes beginning from the perspective of phenomenology, in fact, his construction of phenomenology leaves him open to critiques that his model is observer-centric. I propose that bringing into the conversation other unfamiliar, and thus neglected, but compatible cross-cultural panpsychisms and panentheisms like Abhinavagupta’s, which begin from a phenomenology of subjectivity may be useful for thinking about Tononi’s model.

Particularly Abhinavagupta’s sophisticated and, for his milieu, novel way of linking subjectivity and objectivity in the concepts of knowledge (*jñāna*) and action (*kriyā*)<sup>3</sup> offers a way of understanding how subjectivity can be linked to causality. This particular configuration is mostly absent in Western Cartesian models for understanding consciousness and in Indian philosophical speculations on consciousness. However, this, in any case, is precisely the move that Tononi makes when he proposes that information is both “causal and intrinsic” [2]. Abhinavagupta’s similar linkage of subjectivity with causality can help us to think about Tononi’s neuroscientific mathematical model. For this paper, first I will outline Tononi’s conception of Phi, with some critiques of it. In the final sections, I also address particularly John Searle’s critiques of observer-centrism and Abhinavagupta’s method for linking subjectivity and causality as a model derived from a phenomenology of subjectivity.

## 2. Phi

Tononi’s Integrated Information Theory proposes an information-based theory of consciousness. His theory, drawing from mathematical Systems Theory that models information flows, can be located in a growing body of mathematically based research. Arising out of work begun in the mid-20th century by Claude Shannon and others in the Macy Conferences on Cybernetics ([8], pp. 7, 8),

<sup>2</sup> Searle’s critique of this shift, also asserts that Koch’s subtitle of *Consciousness: Confessions of a Romantic Reductionist* is grossly mis-titled.

<sup>3</sup> The concepts of knowledge (*jñāna*) and action (*kriyā*) are intimately linked to two other terms, *vimarśa* and *prakāśa*. In view of limited space, I do not discuss how these *vimarśa* and *prakāśa* link to knowledge and action in this short paper. I do discuss this elsewhere, (forthcoming).

this notion of information, which paved the way for the development of digital computation, has also been fairly widely adopted by much current research on cognitive science. N. Katherine Hayles has called this the “Information model” of consciousness and has argued that it has fundamentally shifted the way we imagine human consciousness in relation to human bodies ([8], pp. 9–15). Tononi’s model, however, consciously differs from Shannon’s original formulation of information. While Shannon understood information as a reduction of uncertainty through numerical and statistical calculations of probability, Tononi has proposed that we can ascribe meaning and the idea of “experience” to the bits of information that Shannon’s mathematical formulas calculate ([1], pp. 142–45).<sup>4</sup>

What does it mean to ascribe experience to information? With this, Tononi proposes that “there is an identity between phenomenological properties of experience and informational/causal properties of physical systems” [2]. In other words, the information that one can glean from the environment—in the case of the photodiode, an example he uses across his work, information whether the light is on or off corresponds to one integrated bit of information. This simple instrument which responds to its environment by detecting light or the absence of light is able to integrate one bit of information from its two possible concepts, light on or light off, to generate a new irreducible awareness. For Tononi’s IIT this information has material causal effects, and at the same time can be understood as “experience,” with a phenomenological subjective capacity.

The significance of the “experience” of the photodiode is that even this insignificant mechanical detector exhibits a tiny amount of consciousness, though in the case of the photodiode, it is minimal [2]. The photodiode’s new knowledge, one bit of integrated information is what Tononi calls “a difference that makes a difference” [2]. This new bit of integrated information Tononi labels Phi. He proposes the possibility of calculating the amount of Phi in a variety of systems, living and nonliving, as a way of determining the amount of consciousness a system has. Only, however, integrated information, information that cannot be reduced to any smaller component of input, information which offers a “a difference that makes a difference” [2] counts towards Phi. Even though your camera can record far more information than your eye can, nevertheless the camera is not conscious because the information it records is not integrated into new nonreducible knowledge, Phi. So, a human brain, with our vastly expanded neural networks, displays far more Phi than a worm, which displays far more Phi than the photodiode, while the camera on your iPhone has none.

Tononi’s IIT offers five central axioms, which he understands to be starting points, basic things we can say about reality, immediately evident on the face of things. He begins with the first axiom asserting that consciousness exists. With a nod to Descartes’ formulation, he says “I experience therefore I am” [2]. His replacement of “pense,” Descartes’ famous cogito with “experience” not only allows the inclusion of non-humans in the consciousness club, it also aligns the cogito with subjectivity. Because Tononi, as he notes, begins from phenomenology rather than brain neurology ([9], pp. 3, 4, 13, 14)<sup>5</sup>, the subjective sensation of experience, the first-person qualia, is a real and fundamental existent, unlike, for instance, Daniel Dennett’s position, which treats the subjective experience of consciousness as epiphenomenal.

His second axiom is that consciousness is compositional, “each experience consists of multiple aspects in various combinations” [2]. That is, conscious experience is composed of parts; one single experience can have spatial orientation, left and right, a variety of distinct colors, red and blue, all contributing to a unitary experience. His third axiom, Information, asserts that “Consciousness is informative: each experience differs in its particular way from other possible experiences” [2]. Thus, when a person walks into a dark room, the informational component of consciousness entails all the other not actually experienced different possibilities, a light room, a room with a movie playing

<sup>4</sup> Tononi’s whimsical portrayal of the juggling Shannon has him assert, “information is a number, so do not ask it form meaning, nevermind consciousness” ([1], p. 143).

<sup>5</sup> Tononi particularly emphasizes this in his response to Scott Aaronson [9], “Why Scott should stare at a blank wall and reconsider (or, the conscious grid)” (p. 3, 4, 13, 14).

and so on. Tononi points out that the photodiode has only two possible experiences, light on and light off, a vastly diminished repertoire in relation to a person whose differential of information taken from a dark room is exponentially greater.

His fourth axiom, Integration, entails that “each experience is (strongly) irreducible to non-interdependent components” [2]. This is perhaps the heart of IIT, insofar as Phi, the measurement of consciousness, is taken from inputs into a system to produce some new knowledge, which is the integration of the inputs. Tononi gives, by way of example, “seeing a red triangle is irreducible to seeing a triangle but no red color, plus a red patch but no triangle” [2]. The new information, which for Tononi is the experience, is its own irreducible whole and cannot be reduced to its individual components. How do we get from information, i.e., registering whether a light is on or off, to experience? Tononi tells us that concepts are functionally equivalent to *quale*—this is how he makes the bridge between subjective experience (with its unmentioned hints of emotion) to information. Since information can be quantified, he makes it possible to quantify subjective experience [2].

Tononi’s system thus tries to obviate the “hard problem” of consciousness by claiming as axiomatic that, “there is an identity between phenomenological properties of experience and informational/causal properties of physical systems” [2]. This is how Tononi bridges the gap between mind and matter. To bring it to the nitty gritty of how neurons generate this subjectivity, Tononi proposes what he calls the Maximally Irreducible Cause and Effect (MICE) as a way of transposing neuron firing to ideas of subjectivity, such as concepts (apple, table); this he terms “‘*quale*’ *stricto sensu*” [2]. The physical systems formed by neurons make a mechanism that generates information via the brain’s neuronal components; these are understood in Tononi’s system as conceptual structures (“maximally irreducible conceptual structures (MICS)”) which, and this is the lynchpin, he understands as specifying the *qualia* of the experience. In this way he reduces the perennially elusive *qualia* to a composite conglomerate of information. Out of this we get:

An experience is thus an intrinsic property of a complex of mechanisms in a state. In other words, the maximally irreducible conceptual structure specified by a complex exists intrinsically (from its own intrinsic perspective), without the need for an external observer [2].

When Tononi talks about an intrinsic property of a complex of mechanisms he is in effect claiming subjectivity—based on the information generated—for neurons in the brain. We typically do not think of the information generated by our various neurons firing as so intimately productive of the familiar subjective sense “I am here, experiencing this symphonic music” and so on. More typically, we think of the brain’s neurons simply as mechanisms firing on and off but not each contributing to generating our experience when hearing music. To bring this back to his fourth axiom, of integration, Tononi also says that that mechanisms that do not generate integrated information, do not exist from the “intrinsic” (i.e., subjective) perspective of the system [2]. So it is only when information is integrated that you start to get subjectivity.

For Tononi to use words like “cause” and “effect” while at the same time asserting the “intrinsic” nature (i.e., read “subjectivity”) of the mechanism of neurons firing allows him to make a stronger claim than the more commonly pursued correlation claim between neuronal firing and subjective states.

He also points out that one can have a highly complex networked system, a computerized feed-forward system that can win games like “Jeopardy” against humans and can be programmed to recognize natural categories like cats, dogs and pedestrians; however, even such a highly networked system will still be functionally a “zombie” without the step of integrated information. For this reason he cautions against a behaviorist model, based solely on input-output criteria for predicting consciousness [2]. On the other hand, even an extremely simple system, so long as it generates integrated information, i.e., new irreducible knowledge, as the humble photodiode does, can lay claim to consciousness and the possibility of experience [2].

The fifth, final axiom proposes that “Consciousness is exclusive: each experience excludes all others—at any given time there is only one experience having its full content, rather than a

superposition of multiple partial experiences" [2]. This axiom fulfills a functional need of the system. It keeps at bay some unseemly possibilities, for instance, that we might have several neuronal complexes each with their own consciousness in a massive multiple personality mess, or that it might be possible to construe a networked city as having its own consciousness apart from its inhabitants.

So the question is, do we buy it? What's important about the move Tononi makes with integration is that it is precisely a sort of have your cake of materiality and eat it too as subjective qualia and experience. Is this a legitimate move? Is it really the case that if we grant information an "irreducibility" via a complexity arrived through integrating its constituents into a singular concept that this in fact constitutes experience? Is subjective experience really identical with integrated information, as he claims?

Not everyone thinks so. Scott Aaronson in a well-commented blog post critiqued Tononi's theory on the grounds that one could devise a 2-D expander grid to meet the specifications of high Phi, which would necessitate understanding this mechanical device as conscious, and of a fairly high order of consciousness if the grid could be made large enough [10]. Intuitively, Aaronson argues, we need to reject a theory that supposes a 2-D expander grid might be conscious. Tononi's lengthy response, titled, "Why Scott should stare at a blank wall and reconsider (or, the conscious grid)," [11] argues against our "natural" or "intuitive" perspectives on what should or should not count as "conscious," leaving the field open to the likes of well-integrated expander grids, if not IBM's Watson or Siri on your iPhone.

In a different vein, and as a way of segueing into how Tononi's theory of Phi fares in relation to Abhinavagupta's method for integrating subjectivity with causality, I want to point out one other critique of Tononi's theory, one almost reminiscent of Indian philosophical debates on the nature of consciousness. Maguire et al. point out that integrated information entails that some portion of the information is lost; with this loss is the destruction of information, which suggests that Tononi's model does not work ([12], p. 3). What we know about memory is that reaccessing it often strengthens the memory, but if integrating means discarding the independent functions of information, then each time we remember something, some information is lost. So Maguire et al. propose instead that we understand that data compression is what actually happens in Integrated Information Theory. However, following this correction to Tononi's model, they conclude by pointing out that an integrated system which does not have loss cannot be modelled computationally. Thus consciousness in this view cannot be ultimately measured. Given the ubiquitous preoccupation with the problem of memory for discussing the Hindu and Buddhist theories of self and no-self theories (ātmaṁvāda and anātmaṁvāda), the problem of where to store memories, explored by a variety of medieval philosophical schools, is a fitting way to move into Abhinavagupta's understanding of subjectivity and consciousness.

### 3. Abhinavagupta on Linking Subjectivity to Causality

When we think of consciousness, we invariably connect it to sentience, to something being alive. With this, we also popularly connect consciousness with free will, with movement, deliberation and operational causality in the physical world. By way of pointing out the historical contingency of our current position, one might look to classical Indian philosophy, which in many of its philosophical schools reverses this assumption and says that what is truly conscious and sentient is what is precisely without action or causality in the world. Classical Sāṁkhya takes this position when it argues that consciousness, sentience (puruṣa) is like a lame man sitting on the shoulders of a blind person ([13], p. 266). Consciousness, puruṣa, is the essence of subjectivity, but it has no capacity for physical causality in the world. The blind person is the outwardly visible external world, prakṛti, which moves and operates according to physical causal schemes. Similarly, in his famous commentary on the Yoga Sūtras, the 7th-century thinker Śaṅkara tells us that the self (ātman), the nature of which is consciousness and existence, is the knower (jñātr). This conscious self does not perform any action ([14], p. 66), and thus has no causal efficacy in the world. Incidentally, the epiphenomenalism subscribed to by Dennett and others is not really so far off from Śaṅkara's and Sāṁkhya's conceptions



of puruṣa as fundamentally incapable of causing things. The difference is that Sāṃkhya and Śaṅkara understand phenomenal consciousness, puruṣa, as the truly real, while Dennett identifies matter (prakṛti for these systems) as that which is truly real. And of course, Sāṃkhya unlike Śaṅkara maintains the bona fide existential reality of prakṛti.

The 11th-century monistic Kashmiri philosopher Abhinavagupta offers a somewhat different and novel formulation when he argues for matter as fundamentally real. Although the concept of information plays no role in his schema, still, like Tononi, Abhinavagupta proposes a connection between consciousness and matter that asserts their identity. Abhinavagupta thus shares with Tononi a fundamental understanding that links materiality and subjectivity. This, of course, is connected to the practical component of his philosophy, the vyavahāra sādhana ([15], p. 173), a philosophy that understands the world as real and causally operative.

As I pointed out earlier, when Tononi says that information is both “causal and intrinsic” [2] he is making a move similar to Abhinavagupta’s vyavahāra sādhana, understanding the material functions of information, in his case, neuronal firing and computer logic gates, as synonymous with subjectivity in a type of dual-aspect monism. Again, this allows him to make a stronger claim than the more commonly pursued correlation claim between neuronal firing and subjective states.

In a critique of Tononi’s position philosopher John Searle argues out that what Tononi understands as the subjectivity of the photodiode is observer-relative; information itself, including that of the photodiode, is in Searle’s view fundamentally observer-relative ([5], p. 58). What Searle suggests with this is that the information that the photodiode registers, whether the light is on or off, only exists relative to a conscious observer—us. Like the mercury thermometer, the information that is the heat measurement is something measured by us as observers ([5], p. 58). So, while Tononi proposes to begin with phenomenology, Searle would suggest that it is only the phenomenology of conscious human (and animal) observers that counts ([5], p. 58). It can be argued that Tononi’s phenomenology is simply observer-based, not a phenomenology of photodiodes, which in any case may be quite difficult for us humans to obtain, particularly since photodiodes do not have the capacity to express their subjective experiences in forms familiar to us humans. In this sense Tononi’s phenomenology does not quite begin from subjectivity, but rather from the third-person observations of changes noted in photodiode recordings.

Abhinavagupta’s formulation of the link between consciousness and materiality, however, starts unequivocally from a position of subjectivity, which I will delineate in greater detail below. It may be that for Tononi to claim causality and an intrinsic, that is, subjective, component for information, he may need to begin with a phenomenology more strongly centered on subjectivity. Abhinavagupta shows one way of doing this that still affords claims about the reality of phenomena, which is a crucial feature for Tononi’s model, and for that matter, any contemporary model of consciousness that is not simply a form of idealism or dualism.

One difficulty to keep in mind, which may or may not be a stumbling block for Tononi (given already his explicit espousal of panpsychism) is Abhinavagupta’s monism. Panpsychism lends itself naturally to a monism. Our culture’s current dominant model of materialism is also a monism, since there is only simply one kind of stuff in the universe: the building blocks of matter. Our current model is just simply a monism that struggles with what to do with consciousness, since consciousness does not easily fit in with definitions of matter. Chalmers and Nagel in their respective panpsychisms both espouse a kind of dual-aspect monism, two kinds of stuff that are ultimately one. Tononi’s IIT also follows this model.

Abhinavagupta’s monism, which might also be argued for as a dual-aspect monism,<sup>6</sup> begins from the position of subjectivity, rather than observations or phenomena. That is, for Abhinavagupta,

---

<sup>6</sup> Which, by the way, might also be argued as a dual-aspect monism since he uses iterations of knowledge and action; prakāśa (light) and vimarśa (reflective awareness, inwardly focused and outwardly focused as dual components of his monism to afford the subjective and external, causal elements).

consciousness as subjectivity is the basis to describe consciousness, instead of trying to derive the familiar sort of consciousness that includes qualia like the experience of the color red, as Tononi does, from neural connectivity. Thus, rather than trying to build a sense of subjective experience out of neurons, instead the task for Abhinavagupta is to explain how one's subjective experience expands outward into the objective world (presumably into neurons if he had been aware of them).

In a move that mirrors Tononi's assertion that information is both "causal and intrinsic," the third book of the *Verses on Recognition* begins by asserting that action (kriyā), what we normally think of as the operative effects of causality in the external world, is in fact a fundamental component of subjectivity:

In this way, kriyā, Activity, functions both internally and externally following a temporal sequence. It belongs to the Subject alone. Knowledge and actions are inseparably mutually associated with that [Subject]. || 1 || ([16], p. 257).<sup>7</sup>

Here, the subject is not simply the puruṣa or Śaṅkara's knower that we saw earlier, who have no way of affecting the world, no causal efficacy. Rather, this subjectivity encompasses both knowledge as the form of inner awareness and action, which references externality and causality. To reiterate that causal and external properties are always co-existent with subjectivity, in his commentary Abhinavagupta tells us, "in so far in fact as the Sūtra text says 'occurring inwardly and outwardly' it points to action which permeates all through, inwardly also ([16], p. 261).<sup>8</sup> While Tononi needs to assert an "intrinsic" subjectivity for an already obviously causally effective photodiode, Abhinavagupta needs to assert a capacity for causal efficacy by an already assumed subjectivity. Abhinavagupta also explains in his commentary how it is that inner subjectivity is able to encompass the object as an external reality. He says,

The Subject's form, which is a unity of consciousness, contains an excess, an abundance of awareness. This is deposited into the side of the object that is going to be created. So inwardly the [object] has the attribute of śakti, Energy which is none other than the form of consciousness. This is what he points out with the words "having the nature of consciousness" ([16], p. 258).<sup>9</sup>

In effect this argues for a subjectivity, a consciousness for objects, even for photodiodes. The object derives its subjectivity from the subject's form, a "unity of consciousness." One might be inclined to suggest that since it is the subject's form that puts its excess of awareness into the object, this could be construed as a form of observer-centric projection, where a conscious subject, John Searle perhaps, observing the photodiode then deposits the excess of his own awareness into the photodiode, simply imagining that it is conscious, has its own subjective experience based upon its readings of light. To some degree, Abhinavagupta, and his predecessor Utpaladeva, who authored the root text Abhinavagupta comments on in the quote above, seem to support this reading. For instance, elsewhere Utpaladeva notes, "There are two kinds of reality: insentient and sentient. The establishment of an insentient nature rests on the living being" ([15], p. 87).<sup>10</sup> In this respect, insentient objects are derivative; they depend on the subjectivity of a living being for their existence.

In any case, even though objects derive from subjectivity, objects are nevertheless unequivocally real. This is so whether they are in the inward imagination, in the mind of a cognizing subject or whether they have an external existence in the world. Utpaladeva notes, however, that objects have

<sup>7</sup> "evam antarbahir vrttiḥ kriyā kālakramānuḡā | mātūr eva tad anyonyāvīyukte jñāna karmaṇī ||" All translations from the Sanskrit are my own unless otherwise noted.

<sup>8</sup> "tāvad eva āntaratvapratipādanena sarvābhyantarātā api ākṣiptaiva."

<sup>9</sup> Bodhaikarūpasya pramātūr yad adhikamabodharūpaṃ, tat nirmātavyapakṣe niḡṣipyate iti tasya yat nijam śaktirūpaṃ dharmāḥ iti vyapadeśyam, tat bodharūpaṃ.

<sup>10</sup> In Abhinavagupta's commentary on this verse in the *Īśvara Pratyabhijñā Vivṛti Vimarśinī*, vol.1, p. 88, he tells us that objects like blue and yellow reside within the subject, which is the self that is the lord, *Īśvara*, who is eternally perfected.

causal efficacy in the world only when they are externally manifest, not when they exist internally in the mind ([17], p. 114; verse 1.8.6).

However, this causal efficacy always comes from subjectivity, not from an insentient object like the photodiode. Abhinavagupta tells us in no uncertain terms,

it cannot be proved that the relation of cause and effect is grounded within what lacks sentience.<sup>11</sup> Rather, however, the relation of cause and effect lies only with that which is conscious ([16], p. 257).<sup>12</sup>

This is key for maintaining the system's basic proposition that subjectivity can coexist with causality. Also mitigating against a reading that suggests the consciousness of the photodiode is a projection by a conscious observer is that Abhinavagupta tells us in the quote above that "inwardly" the object has its own attribute of "consciousness." This probably bears repeating: the object itself has the nature of consciousness, even in the case where the object, a photodiode or an iPhone, is insentient. Moreover, since Abhinavagupta's system is a monism of subjectivity, the subjectivity of the object, the photodiode, would be accessible to any sufficiently nonlimited consciousness. In practical terms, this might not be the scientist observing the photodiode's on/off responses to light or John Searle, but instead a yogi who has developed a skillful awareness of her own subjectivity to the extent that she can access the subjectivity of the photodiode ([16], p. 311).<sup>13</sup>

The relationship between the inner subjective awareness and a capacity for causal efficacy in the world goes back to Abhinavagupta's dual-aspect monism. They are both the same causally efficacious energy (*śakti*); the former is inwardly directed and the latter is outwardly directed. Abhinavagupta tells us,

When that Energy manifests inwardly, naturally it is called the Energy of Knowledge. However when it expands in stages with its active awareness gradually becoming more firm and fixed, then it manifests externally. This is pointed to as the Energy of Activity ([16], p. 262).<sup>14</sup>

So, the same single Energy (*śakti*) can move in either of two directions, inwardly or outwardly. If it moves inward then it references a subjectivity which is awareness, knowledge. This is what we normally think of as consciousness. Abhinavagupta also tells us that this same consciousness can expand outward, and as it gets more determinate, it becomes an external object. As Isabelle Ratie points in relation to how a yogi's creation accesses a decisive external causality [18], the unifying principle of this Energy allows Abhinavagupta's system to access both an inner subjective awareness and an external worldly causal efficacy at the same time. Experience, the subjective awareness of qualia, seeing the color red, or in Abhinavagupta's example, seeing the color blue, depends upon a unifying activity of consciousness that joins with an object that is something added to it, even as the object exists fundamentally as an offshoot of subjectivity. In an example that calls to mind Tononi's work on deep sleep and persons lacking brain function, Abhinavagupta tells us:

If the host of entities and things, blue etc., are not seen as belonging to the Self, then they would not [appear], but they do appear. When someone says, "this is blue," the person speaking is not devoid of the consciousness which is the Self, i.e., this blue thing is not seen by a fainted person, someone blind or in darkness. Here, seeing blue is necessarily

<sup>11</sup> I.e., mere dead matter –contrary to the assertions of *Sāṃkhya* for example.

<sup>12</sup> "jāde pratiṣṭhitaḥ kārya kāraṇa bhāvo na upapadyate, api tu cidrūpa eva."

<sup>13</sup> "The Subject is called a master when he understands [other objects and entities] to be [merely] the forms of his own limbs." "Svāṅgarūpeṣu bhāveṣu pramātā kathyate patiḥ |". Abhinavagupta discusses these levels of subjectivity at greater length in ([16], pp. 280–320).

<sup>14</sup> Tataḥ sā śaktir antaravabhāsmānasvabhāvatayā jñānaśaktir ucyate, bahirbhāvarūpeṇa tu kramopabr̥ṃhita vimaśadārḍhyena upalakṣitā kriyāśaktir iti vyapadiśyate.



an attribute, something extra added to I-consciousness. Then, having made the two [I-consciousness and the capacity and moment of seeing blue] into a unity, then [we say] a person sees something blue, and there no pure ignorance which is complete lack of perception exists ([16], p. 281).<sup>15</sup>

Abhinavagupta's description of the experience of qualia has suggestive parallels with Tononi's conception of integration. That is, there are two inputs in this schema: here first, the already existing consciousness of the person, as awakesness, not being blind, not being in the dark and second, the color blue. Abhinavagupta says these coalesce, integrate into a single unit, which is the experience, the qualia of seeing blue. Is this akin to Tononi's conception of the irreducibility of consciousness? Abhinavagupta does not, as Tononi does, explicitly focus on the irreducibility of the new experience, the qualia. Moreover, his tendency to subsume experience of the external world into a pervasive, originating subjectivity may make this element of his system less useful or applicable to Tononi's understanding of integration.

#### 4. How to Have a Panpsychism Cake and Eat It Spread with Jam too

Searle has another overarching critique of Tononi's theory, related to the implicit panpsychism it entails. Searle tells us, "consciousness comes in units" and "each consciousness is separate from the other" ([5], p. 57). Poetically, he opines, "[c]onsciousness cannot be spread over the universe like a thin veneer of jam" ([5], p. 57). This critique is in Searle's eyes damning for Tononi's theory. Indeed, can you have a panpsychism (or actually, a panentheism in Abhinavagupta's case), which does not presuppose a unified dispersion of consciousness? Like so much jam spread across all the entities and objects in the space-time of our universe? Tononi's response has to do with consciousness looking more or less like our own depending upon the amount of Phi, integrated information that is irreducible in any given system. Abhinavagupta also addresses this issue, if obliquely. He asks the question, given that everything is ultimately derived from consciousness as an innate subjectivity, how is it that we can get insentient objects at all? In a panpsychist<sup>16</sup> vein, he offers two perspectives, one, we might say, transcendent view which understands everything to be the unfolding nature of consciousness. To this end he says:

However, from the perspective of the manifestation or shining of non-duality, [consciousness] continues without ever coming to an end, even in a corpse, in the body, in a jar and so on. So far as the sphere of non-duality goes, this worldly distinction between sentient and insentient is not made ([16], p. 292).<sup>17</sup>

Everything is ultimately consciousness, (just as in our own contemporary monist materialism, everything is ultimately quarks and particles). The second perspective, keeping in view an awareness of the scaling of the grain, offers levels of consciousness which can be measured. Suggestively paralleling the gradations of consciousness that we find in Tononi's Phi, Abhinavagupta's model offers gradations in grammatical pronomial terms of "Thisness" (idantā) and "I-ness" (ahantā). "Thisness" marks the state of being an object, while "I-ness" marks the state of subjectivity. To the degree that an object participates in "I-ness", subjectivity, it approaches higher and higher degrees of consciousness. Curiously for us, in his system, humans on the whole rank with a surprisingly low amount of his Phi equivalent of "I-ness" ([16], pp. 320–32).<sup>18</sup>

<sup>15</sup> ātmatayāyo na dṛśyate bhāvavargo nīlādih, sa naiva prakāśeta, prakāśate ca asau. Nīlam idam iti hi na mūrccān dhatamasapadamadaḥ. tad avaśyam aham bhāvena adhikena atra bhavitavyaṃ yena kevalam kṛtvā śuddhā tatra ajñatā na bhavati.

<sup>16</sup> Or perhaps panentheist vein. Abhinavagupta is not quite the panpsychist that his great-grand-guru Somānanda is.

<sup>17</sup> Abhedaprakāśanaṃ tu anastamitam eva śavaśarīraghaṭādav api, iti na tatkrto laukiko >yam jaḍajāḍavibhāgaḥ iti yāvat.

<sup>18</sup> Here Abhinavagupta discusses different levels of beings, mostly likely not like us living humans, including gods, *mantras*, *mantramaheśvaras*, *vijñānakevalins*, and so on, who display greater capacities for subjectivity than we mostly do.

He also offers a second answer to the question as well, one that looks a little like McGinn's argument that "the mind-body problem is cognitively closed to humans in the same way that quantum mechanics is closed to a zebra" ([12], p. 5). Pointing to inherent limitations in human and jar awarenesses, Abhinavagupta says,

it is said to be insentient for another reason. This is because [in that state<sup>19</sup>] there is no instrument, no sense organ functioning, which is necessary in order to form a connection with some other thing, separate [from the Subject], and this has to occur in the process of perception everywhere in the world. However this lack of a connection, via the absence of the instrument of perception, the sense organ, occurs not only in Emptiness, etc. It happens in this way, [with a lack of relation] even in a jar. So he says, "hence, in this way" concluding with, "as based upon another." He connects the purpose behind this discussion with our original topic, saying, [this limitation into degrees of consciousness down to the apparently insentient] "derives from vidyā, the limitation of knowledge and kalā, limitation of activity" ([16], p. 292).<sup>20</sup>

Even with a pervasive substrata of consciousness as subjectivity, the particular formations of consciousness as it becomes externalized allow for objects that are insentient, looking like they lack consciousness. So a panpsychism doesn't have to look like a thin veneer of jam. It can look like mountains and humans and photodiodes in this model and still all derive from a basic subjectivity. The differential of apparent consciousness, its seeming encapsulation into units may according to Abhinavagupta's view simply be a product of the limitations in greater or lesser degrees of subjectivity.

## 5. Conclusions

Tononi's model of consciousness shares with Abhinavagupta's fundamental assertion that consciousness and matter are linked through an identity thesis. Tononi links the two through understanding information as containing both subjectivity ("intrinsic property") and causal efficacy; Abhinavagupta does the same, except he uses subjectivity as the basis rather than an idea of information. Given that Tononi's use of information is critiqued by Searle and others as being observer-centric, formulations of the kind of subjectivity-based link between consciousness and materiality that Abhinavagupta espouses present some helpful structural conceptions about how consciousness interfaces with physical causality. In this paper I have not at all focused on Abhinavagupta's theological assumptions, but instead simply addressed conceptual configurations of subjectivity and external causality. What can we take away from Abhinavagupta's model, particularly, what might a neuroscientist gain from paying attention to Abhinavagupta's use of consciousness as the starting point for understanding the world? Abhinavagupta offers a way of resolving the many difficulties plaguing current models of mind precisely through proposing subjectivity as the starting point. Certainly, given the current dominance of a strictly materialist model, Abhinavagupta's proposals to start from subjectivity will likely meet with resistance, in much the same way that the Dalai Lama's calls for first-person science tend to be sidelined. The merit of Abhinavagupta's approach lies in that it puts a theory like Tononi's on a much more solid philosophical grounding.

**Conflicts of Interest:** The author declares no conflict of interest.

<sup>19</sup> Here, he is referring to the state of Emptiness, *śūnyatā*.

<sup>20</sup> Nanu anena prakāreṇa hetunā jaḍamucyate sambandhaḥ parasya bhinnasya sato yat saṃvedanaṃ kartavyaṃ lokayātrothāpakam tasya akaraṇāt. Na kevalam śūnyādāv evam, ghaṭādāv api evam ity āha "ata eva" ityādi "parādhīnatayā" ity antam. Yad artham iyat vicāritaṃ, tat prakṛtam anubadhnāti "vidyākālābhyām" iti.

## References

1. Giulio Tononi. *Phi: A Voyage from the Brain to the Soul*. New York: Pantheon, 2012.
2. Masafumi Oizumi, Larissa Albantakis, and Giulio Tononi. "From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0." *PLoS: Computational Biology* 8 (2014): e1003588. see note 1, above. Available online: <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003588> (accessed on 7 May 2016). [CrossRef] [PubMed]
3. Christof Koch. *Consciousness: Confessions of a Romantic Reductionist*. Boston: MIT Press, 2012.
4. Christof Koch. *The Quest for Consciousness: A Neurobiological Approach*. Denver: Roberts and Co., 2004.
5. John Searle. "Can Information Theory Explain Consciousness?" *New York Review of Books* 10 (2013): 53–58.
6. Thomas Nagel. *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False*. New York: Oxford University Press, 2012.
7. David Chalmers. *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996.
8. N. Katherine Hayles. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature and Informatics*. Chicago: University of Chicago Press, 1999.
9. Giulio Tononi. "Why Scott Should Stare at a Blank Wall and Reconsider (or, the Conscious Grid)." *Shtetl-Optimized: The Blog of Scott Aaronson*. Available online: <http://www.scottaaronson.com/blog/?p=1823> (accessed on 13 April 2016).
10. Scott Aaronson. "Why I Am Not an Integrated Information Theorist (or, The Unconscious Expander)." *Shtetl-Optimized: The Blog of Scott Aaronson*. Available online: <http://www.scottaaronson.com/blog/?p=1799> (accessed on 13 April 2016).
11. Scott Aaronson. "Giulio Tononi and Me: A Phi-nal Exchange." *Shtetl-Optimized: The Blog of Scott Aaronson*. Available online: <http://www.scottaaronson.com/blog/?p=1823> (accessed on 13 April 2016).
12. Phil Maguire, Phillippe Moser, Rebecca Maguire, and Virgil Griffith. "Is Consciousness Computable? Quantifying Integrated Information Using Algorithmic Information Theory." *Cornell University Library*. 1 May 2014. Available online: <http://arxiv.org/abs/1405.0126v1> (accessed on 18 April 2016).
13. Gerald Larson. "Appendix C: The Sāṃkhya Kārikā of Īśvarakṛṣṇa." In *Classical Sāṃkhya*. Delhi: Motilal Banarsidass, 1969.
14. Trevor Leggett. *The Complete Commentary by Śaṅkara on the Yoga Sūtra-s: A Full Translation of the Newly Discovered Text*. New York: Kegan Paul International, 1990.
15. Raffaele Torella. *The Īśvarapratyabhijñākārikā of Utpaladeva with the Author's Vṛtti: Critical Edition and Annotated Translation*. Delhi: Motilal Banarsidass, 2002.
16. Abhinavagupta. *Īśvarapratyabhijñāvivṛtivimarśinī*. Edited by Paṇḍit Madhusudan Kaul Shāstrī. Delhi: Akay Reprints, 1985, vol. 3.
17. Kanti Candra Pandey. *Bhāskarī [Part Three]: An English Translation of the Īśvarapratyabhijñāvivṛtivimarśinī in the Light of the Bhāskarī*. Varanasi: Sampurnanand Sanskrit University, 1998.
18. Isabelle Ratie. "The Dreamer and the Yogin: on the Relationship between Buddhist and Śaiva Idealisms." *Bulletin of SOAS* 73 (2010): 437–78. [CrossRef]



© 2016 by the author; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).