

The integrated structure of consciousness: phenomenal content, subjective attitude, and noetic complex

Katsunori Miyahara 1,2 D · Olaf Witkowski 3,4

Published online: 14 December 2018 © Springer Nature B.V. 2018

Abstract

We explore the integrated structure (or the unity) of consciousness by examining the "phenomenological axioms" of the "integrated information theory of consciousness (IIT)" from the perspective of Husserlian phenomenology. After clarifying the notion of phenomenological axioms by drawing on resources from Edmund Husserl and Maurice Merleau-Ponty (Section 1), we develop a critique of the *integration* axiom by drawing on phenomenological analyses developed by Aron Gurwitsch and Merleau-Ponty (Section 2 & 3). This axiom is ambiguous. It can be read either atomistically as claiming that the phenomenal content of conscious experience is an integrated complex and holistically as claiming that it is an integrated Gestalt. We argue that the latter reading provides a better characterization of the internal structure of the phenomenal content. Furthermore, the integrated structure of consciousness is not confined to the phenomenal content, but it also extends into the subjective attitude (Section 4). Subjective attitudes and phenomenal contents are interdependent constituents that jointly make up conscious experiences. This implies a novel theoretical challenge to the scientific component of IIT, which is to explain how to accommodate the subjective dimension of consciousness into its explanatory scope (Section 5). IIT can respond in a few different ways, but most importantly, it cannot just ignore it once and for all. As one possible way to address the challenge, we propose introducing a novel construct, noetic complex, to develop a fine-grained model of the neural underpinning of consciousness (Section 6).

Keywords Unity of consciousness \cdot Integrated information theory of consciousness \cdot Gestalts \cdot Holistic integration \cdot Aron Gurwitsch \cdot Maurice Merleau-Ponty \cdot Dynamical systems theory \cdot Phenomenology and neuroscience

Neuroscientific theories of consciousness aim to explain the neural substrate of conscious experience. In the early days, the primary approach in the field was

Katsunori Miyahara kmiyahara@rikkyo.ac.jp; miyahara@uow.edu.au

Extended author information available on the last page of the article



experimental, focused on the question of identifying the neural correlates of consciousness in particular brain areas. More recently, however, researchers have begun to pursue more general characterizations of the neurobiological structures that underpin conscious experience (Boly et al. 2013).

The *integrated information theory* (IIT) of consciousness, proposed and defended by the neuroscientist Giulio Tononi and his collaborators, is arguably one of the most innovative and influential among such general theories of consciousness (Oizumi et al. 2014; Tononi 2004, 2008, 2012, 2015; Tononi and Koch 2015; Tononi et al. 2016). The basic idea of the theory is that "consciousness corresponds to the capacity to integrate information" (Tononi 2004, p. 6). We enjoy conscious experience not just because a large number of neurons activate simultaneously or even synchronously, but rather because these neurons interact with one another so densely as to constitute a single integrated unit. More specifically, according to IIT, each conscious experience is underpinned by a set of physical elements with the locally highest value of *integrated information* or Φ , a value that represents the degree of differentiation and integration of an interacting set of elements, which can be measured exactly in mathematical terms.

Proponents of IIT often present the theory as consisting of three components, namely, phenomenological axioms, ontological postulates, and a set of identity claims (e.g. Tononi 2012, p. 296). *Phenomenological axioms* are claims about the essential properties of conscious experience as such. *Ontological postulates* concern the essential properties of the physical substrate of conscious experience. The identity claims, sometimes called *the central identity of IIT* (Tononi 2015; Tononi and Koch 2015), address the relationship between conscious experience and its physical substrate (or more precisely, the abstract, mathematical structure of the substrate).

Our primary aim in this paper is to develop a critical assessment of the phenomenological axioms of IIT. It is a distinctive methodological feature of IIT that it starts from an explicit phenomenological investigation into the essential properties of conscious experience. In our view, however, the existing literature on IIT fails to pay sufficient attention to this aspect of the theory. In particular, we find it unfortunate and surprising that there has been no attempt so far to assess the axioms from the perspective of philosophical phenomenology, that is, the philosophical method of analysis initiated by Edmund Husserl, and developed by a number of philosophers, including Aron Gurwitsch and Maurice Merleau-Ponty, despite IIT's explicit emphasis on the significance of phenomenology. The following discussion will attempt to fill this lacuna by developing a phenomenological critique of the phenomenological axioms of IIT.

This means that the following critique has a limited scope. First of all, one of the main attractions of IIT is the concept of integrated information (or Φ) with the innovative mathematical method to measure it. Given our focus on the phenomenological axioms, however, the mathematical content of IIT lies beyond the scope of this paper. Secondly, some theorists, probably including Tononi himself, consider IIT as entailing some version of panpsychism, according to which everything should enjoy some degree of conscious experience. The argument behind this is that (a) any entity with some capacity to integrate information should instantiate some degree of

¹ For commentaries on the mathematical components of IIT, see for example Barrett (2014), Beaton and Aleksander (2012), de Barros et al. (2016), Peressini (2013), and Tegmark (2016).



conscious experience, and (b) everything has some capacity to integrate information according to the mathematical method proposed in IIT. The first premise (a) is closely related to the central identity of IIT, which we will briefly discuss in the paper. However, we will not offer any explicit discussion on panpsychism in the following.²

Neither will our critique cover the phenomenological axioms in their entirety. In its most recent version, IIT advances five axioms about the essential properties of consciousness under the title of *intrinsic existence*, *composition*, *information*, *integration*, and *exclusion* (Tononi 2015; Tononi and Koch 2015).³ Since our primary interest is in elucidating the integrated structure of conscious experience, our discussion will mostly concern *integration*; otherwise, it will only briefly address *intrinsic existence*.⁴

We believe that the scope of the critique potentially extends to current philosophical discussions on the unity of consciousness (e.g. Bayne and Chalmers 2003; Bayne 2010; Wiese 2017). In particular, our discussion directly concerns what Wanja Wiese calls "the problem of providing a *phenomenological characterization* of R_{PU} [phenomenal unity]" (Wiese 2017, 818). While many philosophers discuss the phenomenon by simply pointing to the fact that multiple phenomenal properties *come together* in conscious experience, we will attempt to clarify what it means from the subjective perspective to have an integrated or unified experience.⁵ In the following, however, we will leave the potential connections to current philosophical debates in the background.

After presenting a preliminary clarification of the notion of phenomenological axioms in section 1, we will develop our critique of *integration* in the next two sections. Section 2 points out an ambiguity of the axiom in its current formulation. Section 3 defends what we call the holistic reading as being more congruent with phenomenological evidence. Section 4 draws out an implication of the holistic reading of *integration*. We will argue that the integrated structure of consciousness is not confined to its phenomenal content, but that it extends into its subjective dimension.

⁵ According to Wiese, investigations from the subjective perspective only provide *possible* solutions to the problem of providing a phenomenological characterization of phenomenal unity. We can determine the *actual* phenomenological character of a unified experience only by exploring the properties of the subpersonal process that realizes the personal level experience (Wiese 2017, 819). In contrast, we think that the phenomenological characterization should be obtained primarily through phenomenological investigations, which can be, but need not be, informed by subpersonal level investigations.



² See Schwitzgebel (2015) for a criticism of IIT in regard to its association to panpsychism.

³ Hereafter, we will italicize the word when it stands for the name of an axiom. For example, "integration" refers to the axiom, and "integration" describes the relevant concept. We will also italicize a word to emphasize it, but then our intention should be clear from the context.

⁴ Just to add a few comments on the other axioms we will not address in the following: The *composition* axiom holds that conscious experience has an internal structure. More specifically, it claims that it is composed of what Tononi calls (not unambiguously) "phenomenological distinctions". The *information* axiom notes that each conscious experience has a specific phenomenal content. This means that each conscious experience has an informational value by virtue of instantiating one content, among the totality of possible contents it can instantiate. The *exclusion* axiom describes the idea that the phenomenal content of a conscious experience is exactly identical with what one is consciously aware of in the experience in question. It forbids us from thinking that there might be more to the phenomenal content than one is consciously aware of, or that only part of what one is consciously aware of constitutes the phenomenal content. These axioms invite philosophical questions no less than *integration* and *intrinsic existence*. For example, what is a phenomenological distinction? What kind of part-whole relation does it have with conscious experience as a whole? (*Composition*) Is informational value an essential property intrinsic to experience itself? Or is it something that the theorist attributes to it from the outside? (*Information*) Can we always draw a definite line between what one is consciously aware of and what one is not? (*Exclusion*) We will not attempt to resolve these issues in this paper, however. See Cerullo (2015) for a criticism of IIT focused on *exclusion*.

The following two sections explore the implication of our phenomenological analysis of the integrated structure of consciousness for the other theoretical components of IIT. Section 5 argues that it poses a novel challenge to the theory, which may or may not be met, but in any case, cannot be ignored. Section 6 outlines one particular way for IIT to answer the challenge by introducing a novel construct called the *noetic complex*.

1 Preliminary clarifications of phenomenological axioms

Tononi insists that scientific theories of consciousness should start from a phenomenological analysis of conscious experience itself. But why should neuroscientists consider phenomenology? Why is it not sufficient to approach the topic objectively, for example, by conducting experiments using behavioral markers of consciousness?

There are at least two grounds on which to answer this question. First, proponents of IIT frequently refer to the heuristic utility of phenomenological analysis. On one account, IIT proceeds by "taking the phenomenology of consciousness as primary, and asking how it can be implemented by physical mechanisms" (Oizumi et al. 2014, p. 2). Phenomenology is useful for scientific inquiry, that is, because it allows the scientists to formulate specific questions that will help them unveil the nature of the physical mechanism underpinning consciousness.

In addition to this, some philosophers argue that phenomenology is not just useful, but even indispensable for any theory of consciousness. According to Taylor Carmon, for example, "[p]henomenology is inescapable, since it is what allows us to specify at the outset what any theory of [...] consciousness must be a theory of (Carman 2005, p. 67, emphasis in original). Conscious experience is by definition a subjective or first-personal phenomenon alien to objective or third-personal observation. We can observe and measure behavioral and neural markers of consciousness, but after all these are only *markers* of consciousness. The usefulness and the necessity of such objective methodologies are undeniable, but assertions about conscious experience are directly confirmed only in first-personal, phenomenological reflections upon conscious experience itself. We can never do without phenomenology, therefore, to evaluate and update theories of consciousness.

There is thus good reason that IIT develops its theory of consciousness based on propositions derived from phenomenological considerations, namely, what it calls "phenomenological axioms" (Tononi 2012, 2015; Oizumi et al. 2014). However, it still needs to be clarified what kind of propositions they are meant to be. The most developed account appears in Tononi's article published in *Scholarpedia*, where he explains that phenomenological axioms are "1. **About experience** itself," "2. **Evident** [in the sense that] they should be immediately given, not requiring derivation or proof," and "3. **Essential** [in the sense that] they should apply to all my experiences." In other

⁷ Michael Cerullo disputes the significance of the phenomenological approach by pointing out that it can lead different theorists to "very different conclusions about fundamental properties of consciousness" (Cerullo 2015, p. 7). This indicates, we suggest, not so much a problem of the phenomenological approach as the difficulty of getting fundamental issues right. In fact, philosophers have arrived at very different conclusions on every fundamental issue throughout the history of philosophy whether or not they preferred the phenomenological approach.



⁶ We thank the anonymous reviewer for pressing us to clarify the role of phenomenology in IIT.

words, they are "meant to capture the essential properties of experience." Furthermore, Tononi states that phenomenological axioms should be "4. **Complete**", "5. **Consistent**", and "6. **Independent**" (Tononi 2015, emphases in original). That is, the axioms should cover all the essential properties of consciousness (4.), they should not contradict with one another (5.), and each axiom should be logically independent from other axioms (6.).

The last three characterizations are not so much about what axioms are as about how they should be formulated. Furthermore, what Tononi means in the third characterization ("Essential") is reasonably clear. It is the first two characterizations, therefore, that we must clarify to better understand what phenomenological axioms are.

What does it mean to say that phenomenological axioms are "about experience itself"? Tononi does not provide a clear answer, but it is possible to clarify the claim by drawing on resources from philosophical phenomenology. Of particular relevance here is Merleau-Ponty's concept of the "phenomenal field" (Merleau-Ponty 2012, p. 52 f.). In our conscious life, we constantly experience a wide variety of objects from our own first-person perspective, including natural things, artifacts, other people, our own mental acts, our own body, the surrounding environment, etc. In short, each segment of conscious experience consists of a totality of appearances of objects in the broad sense, and this totality is what Merleau-Ponty calls the *phenomenal field*.⁸

We are usually more concerned with the objects that appear in the phenomenal field than the phenomenal field itself. However, the latter is not entirely beyond the reach of our epistemic access; we are able to make its contents and structures explicit by reflecting upon our own experience, which is what phenomenology is all about. As Merleau-Ponty puts it, phenomenology is "the study of the *appearance* of being to consciousness" (ibid., p. 62). Drawing on this notion of phenomenology and the object of its investigation, then, we can understand that phenomenological axioms are "about experience itself" in the sense that they are claims about the essential properties of the phenomenal field.

On another interpretation, the axioms are about the "semantic content" of experience (de Barros et al. 2016, p. 57). We doubt that this interpretation does full justice to the claim that axioms are "about experience itself". We also doubt that this is how Tononi considers the axioms. The interpretation in question is probably motivated in part by discussions of "concepts" and "conceptual structures" in IIT (e.g., Tononi 2015; Tononi et al. 2016). The vocabulary may suggest that, for IIT, experience consists in the conceptual content of conscious mental states. However, Tononi applies these concepts neither to conscious experiences as such, nor to the contents of mental state. Instead, he uses them to describe a specific property of the subcomponents of the *physical mechanism* that underpins conscious experience. The relevant property is that of specifying a certain cause-effect repertoire, i.e., a probabilistically distributed set of past and future states causally connected to the current state of the mechanism (e.g., Tononi et al. 2016, p. 452, Tononi 2015, Tononi 2012, p. 301; see also section 5 below). The frequent use of such words like "concept" and "conceptual structure," therefore, does not justify interpreting phenomenological axioms as claims about the semantic or the conceptual content of mental states.



⁸ Not unlike Merleau-Ponty, Elijah Chudnoff attempts to identify the object of phenomenological analysis by calling it "the total phenomenal state" (Chudnoff 2013, p. 560). However, there is a tension between how they each characterize the phenomenal field and the total phenomenal state. Chudnoff stipulates that a phenomenal state is "an instantiation of a determinate phenomenal property" (Chudnoff 2013, p. 561). Hence, for him, the total phenomenal state is a thoroughly determinate domain. In contrast, for Merleau-Ponty, a central feature of the phenomenal field is its indeterminacy. In fact, he considers it a fundamental problem of both psychology and philosophy of his time that they fail to "recognize the indeterminate as a positive phenomenon" (Merleau-Ponty 2012, p. 7).

Next what does it mean to say that phenomenological axioms are "evident", or that "they should be immediately given, not requiring derivation or proof"? It is possible to clarify the claim by drawing on the phenomenological concept of "self-evidence (Evidenz)" (Husserl 2012, p. 284 f.). Consider the Cartesian dictum "I think, therefore I am." Although it is possible to construe this as a conclusion derived from a hidden premise, "Every thinking requires the existence of a thinking self," it is more congruent with the actual epistemic process to consider it as an assertion directly motivated by a reflective observation of one's own experience. By reflecting upon the experience of thinking, that is, we can immediately see that the assertion is true, not unlike how we can immediately see that a physical object exists by observing it in perception.

Husserl describes such experience in which we can directly see the reality of certain objects as bearing *self-evidence* or *insight*; he contrasts it with other types of experience, such as recollection and imagination, where objects appear without motivating us directly to believe in its reality. Drawing on this concept of self-evidence, then, we can understand Tononi as saying that phenomenological axioms are evident in the sense that they concern essential features of consciousness that can be presented with self-evidence by reflecting upon our own experience. In the following, we will express this thought by saying that phenomenological axioms are meant to be supported by *phenomenological evidence*.

Does this mean that there is no room for errors and doubts in the phenomenological axioms? The term "axiom" certainly delivers this impression. In fact, Tononi and his followers do sometimes associate the self-evidence of phenomenological axioms with their infallibility. In one paper, for example, they hold that "axioms are self-evident truths about consciousness [...] that, with Descartes, *cannot be doubted* and do not need proof" (Oizumi et al. 2014, p. 2, emphasis added). In other places, however, it is also suggested that phenomenological axioms are fallible and corrigible. For example, Tononi and Christof Koch proclaim, "Whether the current set of five axioms are truly valid, complete and independent remains open" (Tononi and Koch 2015, p. 5). How can one consider the validity of the axioms to be open if they were truths beyond doubt? It seems that the axioms should be either infallible, incorrigible, and hence beyond doubt, or fallible, corrigible, and hence open to critical assessment.

In which of these two ways should we understand the claim that phenomenological axioms are evident? We take the latter account to be more defensible for two reasons. First, it is more consistent with the actual development of Tononi's thought on the essential features of consciousness. If you see Tononi's exposition of the axioms presented between 2008 and 2015, it is clear that he has constantly updated his view through the course. It is difficult to make sense how this would have been possible if phenomenological axioms were incorrigible.

Second and more importantly, it is more consistent with the actual character of epistemic inquiries to distance self-evidence from infallibility. Assertions are not true without any qualification just because they are based on self-evident experience. Even when you assert the existence of an object based on a perceptual experience of it, for example, it is always possible that the assertion turn out to be false on account of a later experience that reveals the illusory character of the original "perception" (Husserl 2012,

¹¹ We thank the anonymous reviewer for pressing us to clarify this point.



¹⁰ See Husserl (2012), pp. 284–291, for more discussion.

p. 289). Hence, we must not think that a claim "cannot be doubted" just because it is supported by phenomenological evidence. This does not deny that some self-evident claims may be infallible or incorrigible. When it comes to the phenomenological axioms, however, unless there are other stronger reasons to believe otherwise, we should construe them as fallible and corrigible, just as perceptual judgments supported by self-evident observations are.

With these clarifications in place, we can now move onto a critical assessment of the phenomenological axioms as it is currently proposed in IIT.

2 Ambiguity in the integration axiom

In this section and the next, we will develop a phenomenological critique of the *integration* axiom. As you can see from the title of the theory, this axiom has played a central role in the development of IIT since its most early stages—that is, since before Tononi started to make explicit use of the term "phenomenological axioms." Here are three different ways in which he advances the idea.

The integration of information in conscious experience is evident phenomenologically: when you consciously "see" a certain image, that image is experienced as an integrated whole and cannot be subdivided into component images that are experienced independently. (Tononi 2004, p. 3)

Phenomenologically, every experience is an integrated whole, one that means what it means by virtue of being one, and which is experienced from a single point of view. (Tononi 2012, p. 295; see also Tononi 2008, p. 219)

Consciousness is *unified*: each experience is *irreducible* to non-interdependent, disjoint subsets of phenomenal distinctions. (Tononi 2015; see also Tononi and Koch 2015, p. 6)

The idea seems simple and clear: Consciousness is an integrated or a unified whole; it is not just the sum total of its components. Even if we limit our attention to visual experience, for example, it is not just a combination of the left part and the right part of the visual field (Tononi 2004, 2008, 2012, 2015; Tononi and Koch 2015). To see a blue book is not just to experience the blue color and the square shape of the book at the same time (Tononi 2012, 2015; Tononi and Koch 2015). However, there is a hidden ambiguity in this axiom as it is expressed in these formulations. As we will argue in the following sections, furthermore, it is crucial to dissipate this ambiguity in order to be clearer about the precise scope of the integrated structure, which IIT maintains is essential to consciousness.

Before discussing the ambiguity, however, it is important to clarify the subject matter of the axiom—that is, what *integration* is a claim about. Although it is posed as a claim about "conscious experience" or "experience," if you look at what is actually said, it only appears to address what we may call the *phenomenal content* of conscious experience. Phenomenologists in the Husserlian tradition typically consider conscious experience to consist of a *subjective act* and a *phenomenal content*, or a "noetic" and a



"noematic" content in Husserlian terminology (Husserl 2012, p. 182 f.). When you see a blue book, for example, your experience consists of the act of seeing, which is related to the phenomenal content of the experience which includes the book as it is presented to the experience. When you imagine the blue book in your bookshelf without seeing it, you have an experience with a similar phenomenal content, but this is correlated to a different subjective act, namely, the act of imagining. Accordingly, Husserl notes that phenomenological descriptions can be developed in two directions:

[O]n the one hand, [there are] descriptions of the intentional object as such, with regard to the determinations attributed to it in the modes of consciousness concerned, [...] This line of description is called *noematic*. Its counterpart is *noetic* description, which concerns the modes of the *cogito* itself, the modes of consciousness (for example: perception, recollection, retention), with the modal differences inherent in them (for example: differences in clarity and distinctness). (Husserl 1960, p. 36)

From this perspective, any phenomenological description of conscious experience will remain incomplete until it addresses both its noetic and noematic dimensions. In developing the phenomenological axioms, however, Tononi is almost exclusively concerned with the phenomenal or the noematic content, paying much less attention to the subjective or the noetic act. Based on this distinction, you can see that *integration* only holds that the phenomenal content of consciousness is an integrated whole. (We will come back to the issue of the subjective or noetic dimension of consciousness in section 4).

The axiom is still ambiguous, however, as to the precise meaning of integration. In its current formulation, it allows us to read it both atomistically and holistically. The two readings envision the integrated structure of the phenomenal content under different concepts of wholes: The *atomistic reading* takes *integration* as claiming that the phenomenal content is an integrated *complex*; the *holistic reading* takes it as claiming that the phenomenal content is an integrated *Gestalt*.

What is an integrated complex? To see this, it is useful to compare it with another kind of a whole, a sum. Uriah Kriegel explains that their difference lies in the interconnectedness of the parts:

Intuitively, the idea is that complexes involve an essential interconnection among the elements, whereas a sum is but the mere compresence of those elements. Consider the parts of an Ikea table lying about in a box, pre-assemblage, and the same parts put together so as to form a functional desk. The former is a mere sum of the parts, the latter a complex. Accordingly, it is possible to destroy the desk without destroying any of its parts, namely by destroying the interconnection among them; but the sum it is possible to destroy only by destroying one of the parts. More generally: a collection C of parts P1,..., Pn is a sum iff the only way to destroy C is to destroy a Pi; C is a complex iff it is possible to destroy C without destroying any Pi. (Kriegel 2015, p. 162)

A complex is an *integrated* whole in a way that a sum is not because of the interconnection among the parts, which is essential to make the whole the whole it is.



How is a Gestalt different from this? Gestalts and complexes both can be destroyed by just destroying the interconnection among their parts. But they are different in the ways in which each of their composing parts maintain their identity. The composing parts of a complex retain their identity even after the interconnection among them is destroyed; the composing parts of a Gestalt can alter their very identity as a result of the destruction of the interconnections. When it comes to Gestalts, in other words, not only the identity of the whole, but also the identity of the composing parts depend upon the interconnection among them.

To use a different pair of concepts discussed by Aron Gurwitsch, a complex is an integrated whole made up of merely interconnected *elements*, while a Gestalt is an integrated whole made of interdependent *constituents*. Elements "may be experienced both in isolation and in grouping without being internally affected by either mode of appearance. [...] [E]lements preserve their identity whether they are grouped or not" (Gurwitsch 1964, p. 144). In contrast, constituents are what they are in virtue of their "functional significance," that is, the role they play to constitute the whole, and therefore they may fail to retain their identity when isolated from their current context:

If a part is defined by its functional significance and, thus, proves dependent upon and determined by its co-constituents, it is the contexture of the co-constituents that qualifies the part under discussion. [...] Qualified and defined by its functional significance with regard to other constituents, the part exists as that which it is only in its reference to the whole. Such reference is inscribed and included in the nature of the part, not as an additional modification, but as an essential determination. [...] It is through realizing the whole that any part is made to be at its locus what it actually is. (Gurwitsch 1964, p. 145, emphasis in original)

A constituent needs to be in the right context to maintain its identity. What it is is determined in part by the co-existing constituents and how they jointly make up a particular whole. As an illustration, consider team membership (Varzi 2016). In a rugby team, for example, the fly-half is a fly-half only because of her functional role in the whole team, or only in virtue of her relation to the other members of the team who play different positions so that they jointly constitute a team. If you isolate the fly-half from the team so that she cannot join the game, then the team will cease to exist—unless a different player takes her place to play as a fly-half—which shows that the team is not a mere sum of players. In addition, the one isolated from the team will no longer be a fly-half independent of her functional role in the team, which means that she has been a part of the team as a constituent rather than an element. In other words, constituents stand in a relation of "metaphysical dependence" (Chudnoff 2013, p. 564) to the other co-existing parts. In this sense, Gestalts consist of interdependent constituents, while complexes consist of independent, but interconnected elements.

In light of this distinction, we can specify the ambiguity of *integration* as follows: it remains unclear whether it claims that the phenomenal content is an integrated complex or an integrated Gestalt. In fact, the axiom in its current formulation only states that the phenomenal content is *not* a sum total of its composing parts. Then how should we interpret *integration*? In which sense are we supposed to understand that the phenomenal content is an integrated whole?



3 Holistic integration of the phenomenal content

The previous section pointed out an ambiguity involved in the *integration* axiom. This section argues that we have good phenomenological reasons to adopt the holistic reading, according to which the phenomenal content of conscious experience is an integrated Gestalt made up of interdependent constituents.

To see how the holistic reading is better supported by phenomenological evidence, let us first consider Gurwitsch's phenomenological analysis inspired by Gestalt psychology of the experience of seeing an image of Rubin's vase (Gurwitsch 1964, 117–121). You can see the image either as that of a white vase or as that of two faces facing each other (see below). What exactly happens in our experience when we see the image alternate its appearance? In particular, how does the boundary between the white and black areas appear in the course of this experience?:



One might think that the visual content of the experience is constant throughout the alternation: It is only the way we interpret the image that switches from one to the other; in particular, the phenomenal appearance of the boundary between the two areas never changes through the course of experience. Such account, however, proves inconsistent with the phenomenological evidence. For a careful phenomenological reflection reveals how the phenomenal appearance of the boundary itself varies depending on the phenomenal appearance of the other parts of the phenomenal content. When the white area of the image appears as a vase, the black area appears as the background, and the boundary appears as the contour of the vase or as delimiting the white area. Alternatively, when the black area appears as two faces, the white area appears as an opening in the background, and the boundary appears as the outline of the faces, that is, as delimiting the black area. Either way, the phenomenal character of the boundary is determined by the way its co-existing parts appear to us. In other words, they jointly make up the whole phenomenal content as interdependent constituents rather than merely interconnected elements.

¹² Carlos Montemayor and Harry Haladjian, for example, account for the perceptual experience of the duckrabbit image and of the Necker cube in the following way: "Neither of these shifts in perception entail changes in the object itself. The corresponding images, with their static and objective features, would constitute the icon, while that experienced changes would be characterized in terms of semantic content, attention, and epistemic access to mutually incompatible interpretations of the stimuli" (Montemayor and Haladjian 2015, p. 145).



The holistic structure of the phenomenal content is not only found in such experimental settings. Consider the experience of seeing a blue and wooly rug. One might think that the blue color and the wooly texture are integrated in the experience in the sense of being attributed to the same object, the rug, but that each quality is determined independent of one another. As Merleau-Ponty points out, however, in perceptual experience, the color of an object is always co-determined by its other qualities: "A color is never simply a color, but rather the color of a certain object, and the blue of a rug would not be the same blue if it were not a wooly blue" (Merleau-Ponty 2012, 326). The phenomenal appearance of the blue color of the rug is what it is only in virtue of the co-existing phenomenal appearance of the wooly texture. Even if a porcelain plate reflected light with the same wavelength as the blue rug, for example, the phenomenal appearance of their surface color would be entirely different. We can talk about the experience as containing a blue color and a wooly texture as if they are mutually independent qualities, but the phenomenal appearances are interdependent on one another.

The effect may be more manifest in cases that involve a dramatic change in the phenomenal content. Consider the following scenario: you are walking in a forest, and you see a flat, white, smooth stone lying ahead on the ground. Not wanting to slip on it, you prepare yourself to walk cautiously, but then after taking a few more steps forward, you notice that you are just seeing a patch of sunlight cast on that part of the ground (ibid., 310). In this example, your initial experience contained a white color and a smooth texture, while your later experience involved a light brown (of the ground illuminated by sunlight) and its rough texture, but the transformation does not happen part by part. You do not experience the transformation of the color and of the texture independent of one another. The change in color from white to light brown already implies the change in texture from smooth to rough, as well as the change in the identity of the object from a flat three-dimensional stone placed on the ground to a twodimensional patch of light cast on the ground. The phenomenal appearances of the color, the texture, and the object itself are all determined in relation to one another. Once again, they are mutually dependent constituents, rather than merely interconnected elements, of the phenomenal content as an integrated whole.

One might object, however, that these examples only show how the thematic object shows up as an integrated Gestalt. Suppose we are right to say that Rubin's figure, the blue rug, and the patch of sunlight figure in the phenomenal content as integrated Gestalts made up of interdependent constituents. Even so, this does not show that the *entire* phenomenal content (at a particular moment) is an integrated Gestalt. For phenomenal contents are not exhausted by the thematic object; they always include thematic parts and non-thematic parts, parts of which we are explicitly aware and parts of which we are not, parts in the foreground and parts in the background. To argue that the entire phenomenal content is integrated holistically as a Gestalt, we must show that the non-thematic parts are also integrated with the thematic parts as interdependent constituents of the whole.¹³

In the first place, we can respond to this objection by using examples that indicate how the thematic and the non-thematic parts are not just co-existing elements. To draw on Merleau-Ponty once again, when we see an object at a particular distance, the



¹³ We thank the anonymous reviewer for pressing us to address this problem.

experienced distance is determined not only by the appearance of the object in question, but also in part by other objects of which we are not aware thematically:

[O]bjects interposed between me and the one I am focusing upon are not perceived for themselves. But they are, nevertheless, perceived, and we have no reason to deny this marginal perception a role in the vision of the distance since the apparent distance shrinks the moment a screen hides the interposed objects. [...] [T]he bell tower [...] appear to me as smaller and farther away the moment that I can see more clearly the details of the hills and the fields that separate me from it. (Merleau-Ponty 2012, pp. 49-50)

The same effect can be observed, for example, when we see the coastline from the ocean. Since there are only few objects, if any, interposed in-between, the coastline appears much closer to where we are than it really is. Notice that depth is an indispensable dimension of spatial perception: To see an object is always to see it at a distance. Accordingly, these examples strongly suggest how the phenomenal appearance of the thematic content is not just determined by its internal constituents, but that it is always determined in part by the non-thematic parts of the phenomenal content ("marginal perception") as well.

Elijah Chudnoff (2013) makes a different case for a holistic interpretation of the unity of phenomenal content by resorting to a general relationship that holds between the thematic and the non-thematic parts. There is a stark phenomenal contrast between these two parts: the thematic parts conjoin with one another to jointly form a cohesive individual; the non-thematic parts also conjoin with one another, but only to form indefinite units. For example, recall how the black and the white areas in Rubin's figure are not delimited when they appear as the background of what is in the foreground. This phenomenal contrast might tempt one to suspect that the thematic and the non-thematic parts are merely co-existing elements of the whole content.

However, whether a part of the phenomenal content appears thematically or non-thematically depends on its relation to other parts, or more specifically, on its relevance to the thematic content (Gurwitsch 1964, pp. 340–341). Non-thematic parts of the content appear non-thematically precisely because they are less relevant to the thematic parts than the latter are among each other; thematic parts appear thematically precisely because they are highly relevant to one another. In Chudnoff's words, the "phenomenal manner" in which any part of the experience figures in it—whether it appears thematically or non-thematically—is determined in part by its relation to the thematic content. If this is true, phenomenal appearances of the thematic and the non-thematic parts are determined in part by their relation to one another, which is to say that the entire phenomenal content is a holistically integrated Gestalt that consists of interdependent constituents.

Is it not still possible, however, to consider the phenomenal content as nothing more than an integrated complex composed of phenomenal elements? These elements may *appear* to change their quality in relation to their co-existing elements, but this may

¹⁴ What Gurwitsch means by "relevance" is not entirely clear. Chudnoff explains it as a relation of "centrality" and points out how it can be determined in various ways. See Chudnoff (2013), pp. 569–571 for more discussion.



only be a matter of appearance. For example, one might envision conscious experience as consisting of two layers of content, namely, iconic content and semantic content (Montemayor and Haladjian 2015, pp. 145–149). On one definition, the iconic content of a visual experience consists of a "preinterpreted, strictly structural features of a visual scene" (ibid., p. 145); the semantic content consists of the visual scene interpreted in terms of specific types of visual objects.

Take the experience of seeing a Rubin's vase/face figure. The iconic content is something like a visual image that strictly represents the configuration of the black and white areas and nothing more than that. In the semantic content, then, the figure is interpreted as a depiction of a vase or of two faces. Based on this distinction, one might claim that the phenomenal content is exhausted by the iconic content. We argued earlier that the boundary between the two areas changes its appearance depending on the way the rest of the figure appears to us. However, this is not accurate. The phenomenal appearance of the boundary is invariant. It is only the way in which we interpret it that changes through the course of experience. It is entirely possible, one might therefore argue, despite all the examples meant to show the holistic nature of the phenomenal content to defend its atomistic nature: The semantic content is determined holistically; however, the phenomenal content itself is determined through and through by the invariant qualities of its composing elements.

This account of the phenomenal content is not defensible, however, from a strictly phenomenological perspective. It is indeed possible to explain variations in phenomenal appearance by positing an invariant, iconic content beneath them. However, the invariant layer is not something about which we have direct phenomenological evidence. Even to say that the iconic content underlies the varying appearances, we have to admit that what we undergo in our experience is the variation. Phenomenologically speaking, therefore, the phenomenal content is better understood holistically as an integrated Gestalt made up of interdependent constituents, rather than an integrated complex made up of merely interconnected elements.

All in all, there is good phenomenological reason to envision the phenomenal content holistically as an integrated Gestalt made up of interdependent constituents. We conclude for this reason that the holistic reading of the *integration* axiom is preferable to the atomistic reading. But then does the *integration* axiom interpreted under the holistic reading express everything that needs to be said about the integrated structure of consciousness? This is the topic of the next section.

4 Integrating the subjective dimension of consciousness

We submitted that the phenomenal content is integrated in the sense of being an integrated Gestalt. In this section, we continue our phenomenological analysis to argue that the integration of this kind is not confined to the internal structures of phenomenal contents but that it extends into the subjective dimension of consciousness.

The subjective dimension is an aspect of consciousness that is hardly addressed in the phenomenological axioms. The only place they barely mention it is in *intrinsic existence*, which Tononi formulates as follows:



Consciousness *exists*: each experience is *actual*—indeed, that my experience here and now exists (it is real) is the only fact I can be sure of immediately and absolutely. Moreover, my experience exists from its own *intrinsic* perspective, independent of external observers (it is intrinsically real or actual). (Tononi 2015, emphasis in original)

As we understand it, this axiom points to the subjective nature of conscious experience. Your experience of the book has an actual existence just as the book itself does, but their modes of existence are not the same. The mere existence of the book is not sufficient for there to be an experience of the book. For this, the book must be seen. There must be a subjective act of seeing correlated to the book. Consciousness is intrinsically subjective in the sense that it always involves a relation to a phenomenal content, an experiential relation that is also a part of each conscious experience. Accordingly, we can never understand what consciousness is without taking the subjective dimension into consideration. We take it that this is the idea Tononi tries to deliver in *intrinsic existence*. ¹⁵

Note that the claim here is not that consciousness requires a conscious subject, a persistent bearer of conscious states. ¹⁶ To report our current conscious experience, we usually use expressions such as "I see a book" or "I am imagining a unicorn". Based on such expressions, one might think that the presence of a conscious subject, the I, is a necessary component of all conscious experience; furthermore, this may lead one to think that consciousness requires self-consciousness, that is, an explicit or implicit awareness that you are the invariant bearer of the various conscious states that occur in your experience. It is not this line of thought—what Gurwitsch calls the "egological conception of consciousness" (Gurwitsch 1966b, p. 288)—that is at issue when we talk about the subjective nature of consciousness. When we say consciousness has a subjective dimension, we only mean that it involves an experiential relation to the phenomenal content, also known as the *noetic act* or *noesis* in Husserl's terminology.

We think Tononi is right in pointing out that consciousness always involves a subjective perspective. In his writings, however, he does not provide much positive characterization of the subjective dimension essential for conscious experience. But then how exactly are we supposed to understand its relation to the phenomenal content? We will take this question as a guide to further explore the integrated structure of consciousness.

4.1 The basic phenomenological argument

To answer this question, we can start by contrasting two ways to think about the issue. On the one hand, we can think of the subjective dimension as an *impotent perspective* onto the phenomenal content. The phenomenal content is determined in itself, and the only role of the subjective dimension is to bring this ready-made content to conscious

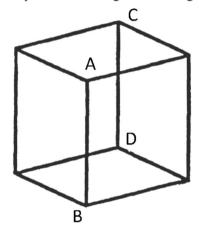
¹⁵ Anthony Peressini advances a different interpretation. He argues based on a careful reading of Tononi's writing that IIT means by "intrinsic" that something is an "observer independent (real) property" (Peressini 2013, 199–200). We admit he is exegetically correct, and so the criticism is effective. But we do not think that this is the idea Tononi actually tried to deliver with this expression. For it is clear from what he writes that he writed to emphasize how consciousness has a different mode of existence from other objective entities.





awareness. On the other hand, we can also think that the subjective dimension is formative of the phenomenal content. Each phenomenal content is what it is partly in virtue of the character of the subjective dimension. Tononi's choice of the word "intrinsic perspective" strongly suggests the impotent conception, while the term "subjective act" is more suggestive of the formative conception. However, IIT is not explicitly committed to either conception of the subjective dimension.

Then which conception of the subjective dimension fits better with the phenomenological evidence? Let us approach this question by first looking into a phenomenological analysis of Necker cube illusions (Ihde 2012, p. 65f; Merleau-Ponty 2012, pp. 273–276). The experience of seeing a Necker cube figure is often characterized in terms of a *spontaneous* alternation between two cube appearances. Don Ihde points out, however, that being spontaneous is not an essential feature of the experience. In fact, there is a particular correlational structure between the cube appearance and the way one sees the figure, and once this is recognized, one can gain some control over the appearance. Consider the experience of seeing the following figure:



The appearance in which line AB is the closest to you correlates with a perspective where you look down at the cube, and the other appearance in which line CD comes forward correlates with a way of seeing where you look up at the cube. The looking down or looking up describes an imaginative direction of seeing, rather than the actual direction that is determined by one's spatial relation to the figure itself. Even though the actual direction is constant, thus, in this experience, your imaginative perspective on the cube varies in tandem with the alternation between the two cube appearances. Based on this insight, then, you can somewhat alter the appearance at your will by altering the subjective act of imaginative seeing at your will. By looking down, thus, you can make the cube with line AB in the front show up, and by looking up, you can make the other cube appear instead. This is not to say, however, that you have complete control over the appearances. Despite your attempt to maintain a specific way of seeing, the figure itself often influences the way you see it, and this may result in an alteration of the appearance against your will. Merleau-Ponty denies, accordingly, that the way in which we see the figure is a matter of free, unconditioned choice. Rather, "[t]hey are indicated or recommended by the phenomena." (ibid., p. 274).



A similar effect can be identified in Rubin's figure discussed in the last section. In this case, it is perhaps easier for most people to change the appearance of the whole figure by trying to see the vase or the two faces. But again, typically, you will not have complete control over the appearances. Even when you see the vase in the foreground and the black area recedes to the background, the latter can somehow make you focus on them, and eventually, you might be led to see them as two faces in the foreground. Then perhaps after a while, the white area may begin to draw your attention, which will once again, lead to a reversal of the visual content.

These examples indicate a dynamic interplay between the subjective dimension and the phenomenal content. The interplay consists in how the two parts of consciousness are determinant of one another. On the one hand, the subjective act is formative of the phenomenal content. What appears in your phenomenal field in seeing the Necker cube is determined in part by how you see it.¹⁷ On the other hand, the phenomenal content is also formative of the subjective act. How you see the Necker cube is determined in part by what appears in your phenomenal field. In other words, the phenomenal content and the subjective dimension interacts with one another in such a way that each part's identity is determined in part by the other part with which it co-exists.

If this is true, it means that the subjective act and the phenomenal content are interdependent constituents of the whole conscious experience. For this reason, we claim that the integrated structure of consciousness is not confined within the phenomenal content, but that it extends into the subjective dimension as well. Each conscious experience, therefore, is an integrated Gestalt made up of a subjective dimension and a phenomenal content. In the rest of this section, we will elaborate on this idea by examining three possible objections.

4.2 How subjective acts form phenomenal contents

The first objection addresses the claim that subjective acts are formative of phenomenal contents. One might say that it plainly goes against phenomenological evidence, particularly, in regard to cases of perceptual experience. We simply do not have the power to determine what appears in our phenomenal field. At least when it comes to perceptual experience, then, should we not prefer the impotent model to the formative model of the subjective dimension?

We can answer this objection by clarifying what we mean by the *formative effect* of the subjective act. By this, we mean to describe how the subjective act pre-disposes the self-organization of the phenomenal content in a certain direction. In the case of the Necker cube, even after you learn how to exercise some voluntary control over the appearance, it takes a short period of time for the transition to take place after you adopt a particular way of seeing. You do not have the intended appearance the moment you adjust your imaginative perspective. It is only after you let the other parts of the figure find their own place and take on a particular shape by themselves that you actually have the intended cube show up. As Merleau-Ponty notes, "Even in cases where the

 $^{^{17}}$ de Barros et al. (2016) indicates a similar point in terms of the "contextuality" of consciousness. They suggest based on this observation that it may be impossible to make exact calculations of Φ -values, not just as a matter of practical difficulty, but as a matter of mathematical incompatibility between IIT and the mathematical theory of contextuality.



organization is ambiguous and where I can make a shift, I do not achieve this directly [...] I am sometimes obliged to wait for the organization to produce itself' (ibid., p. 275). The subjective act contributes to the formation of the phenomenal content by exerting some influence to its self-organization. Understood in this way, the formative conception should appear much more congruent with the phenomenological evidence.

4.3 The role of attention in ordinary experience

The second objection concerns the scope of the argument. One might suspect based on the nature of the example we used in the argument that it only applies to conscious experiences in experimental settings. We can answer this objection by listing more ordinary examples, in particular, by observing the dynamic interplay between attention and the phenomenal content in those cases.

Before moving onto the actual response, however, we should clarify the concept of attention at issue.¹⁸ Recent cognitive scientists, psychologists, and philosophers typically adopt a "function-centered approach" (Wu 2014) to attention, which identifies it with some functional role in the information processing of the cognitive system. In particular, most attention researchers consider the functional role at issue to be some form of information selection, even if their views widely differ as to the precise nature of the selection in question.¹⁹ On this approach, attention can shape cognition and behavior without necessarily affecting conscious experience. Accordingly, in recent attention research, there has been a heated debate about its precise relation to consciousness—in particular, whether attention is necessary for conscious experience (e.g. Prinz 2012) or not (e.g. Koch and Tsuchiya 2007).²⁰

In contrast, we will operate on a phenomenological conception of attention, according to which it is a subjective component of consciousness that plays some functional role *within conscious experience*. As a first approximation, the functional role at issue is that of transforming the phenomenal field. Attention is a process that leads one phase of experience to another. The phenomenological approach to attention is concerned with elucidating the nature of this process, in particular, by asking how it is induced from within the current experience and then how it gives rise to new phenomenal contents. As Merleau-Ponty puts it, the question of attention consists in "show[ing] how a perception awakes attention, and then how attention develops and enriches this perception" (Merleau-Ponty 2012, p. 29).

Since our focus is only on illuminating the relation between the subjective component and the phenomenal content within conscious experience, we will not delve deeply into the phenomenology of attention.²¹ It is important to keep in mind, however, that the following discussion is pitched at the phenomenological level, and that attention is not considered as a type of information processing. Although we are going to discuss the role of attention in consciousness, thus, we will not advocate any particular position in the contemporary debate concerning the relation between the two terms.



¹⁸ We are grateful to the anonymous reviewer for pressing us to be clear on this issue.

¹⁹ See Wu (2014) for more discussion.

²⁰ See Montemayor and Haladjian (2015), for a comprehensive overview and a critical discussion of the debate.

²¹ See Arvidson (2006) for a contemporary discussion on this topic.

With this phenomenological conception in mind, let us consider how attention relates to the phenomenal content in ordinary situations. Take an example discussed by Tononi himself: the experience of lying on an armchair, seeing a blue book in your room. If you shift your attention from the book to the picture placed on top of the bookcase, this will change the phenomenal appearance both of the book and of the picture. The shift of attention decreases the clarity and the distinctiveness of the book and its properties, such as its shape, color, and texture, in your phenomenal field, while conversely increasing those of the picture. Moreover, even though you are usually able to shift your attention at your will and change the ways in which things appear to you accordingly, your attention is not under full control. Once your attention is directed at the picture as a whole, for example, it is not easy to retain this attentional attitude, for typically other objects in the phenomenal field (the details of the picture, other things or other people in the room, etc.) begin to draw your attention before long. Just like the previous examples, in short, the phenomenal content is partly determined by the subjective act of attending, and vice versa. This example suggests that even in ordinary cases, conscious experience consists of a subjective act and a phenomenal content as integrated constituents formative of one

Admittedly, the formative effect of the subjective act is not exactly identical with the previous case. In the case of seeing a Necker cube, changes in the subjective act affect the very identity of the object that figures in the phenomenal content. When your imaginary viewing attitude towards the Necker cube figure changes, you see a different cube. In the current case, in contrast, changes in the subjective act only affect the way in which an object figures in the phenomenal content. When your attention shifts away from the book, and the book recedes to the periphery of your visual field, appearing unclear and indistinct, for example, you still continue to see the same book; it just appears differently.²²

This difference, however, does not affect the claim that the subjective act is formative of the phenomenal content in both experimental and ordinary settings. For two experiences presenting the same object in two different ways are different in regard to their phenomenal contents. Even if the attentional shift had no effect on the identity of the object, therefore, as long as it changed the way in which the object appears in the phenomenal content, the phenomenal content is what it is partly in virtue of the subjective act. If this is true, we should think that the subjective act and the phenomenal content are formative of one another not only in experimental settings, but in ordinary situations as well.

²² Chudnoff describes the effect by making a conceptual distinction between two kinds of phenomenal properties, *phenomenal content* and *phenomenal manner* (Chudnoff 2013, pp. 565–6). Phenomenal content is determined by the things that are presented in experience, while phenomenal manner concerns the way in which they are presented. On this terminology, the phenomenal content changes in the case of Necker cube, while in the ordinary case of attentional shift, it is only the phenomenal manner in which the content shows up that changes. In contrast, we use the term *phenomenal content* to describe the totality of what correlates to the subjective act, including both the things presented and the ways in which they are presented, which roughly corresponds to what Chudnoff calls *phenomenal character*. For the various ways in which conscious experience is modified by attention, see Arvidson (2006), ch.3, and Gurwitsch (1966a), pp. 223–250.



4.4 Subjective acts and mere subjective attitudes

A different line of objection might be raised from reflections upon what is sometimes called absorbed coping (Dreyfus 2005) or the experience of flow (Csikszentmihalyi 1990). Absorbed coping is a form of bodily action that achieves its goal without involving explicit self-awareness of the action itself or of the goal. The most obvious cases of such experience are found in expert performances, for example, by professional musicians or top-level athletes; however, they also show up in ordinary, habitual behaviors such as greeting people, participating in rituals, or even adjusting your distance from other people in socially appropriate ways.

Instead of shaping our own action at our own will, in all these cases, we are completely immersed in the situation that draws the action out of us: "the situation itself presents to me "reasons" for action that immediately draw on my body, soliciting a response" (Wrathall 2014, p. 3). In other words, we are "just doing what we do" (Dow 2017) without being self-aware of what we are doing or even that we are doing what we are doing.²³ This suggests that the experience of absorbed coping is exhausted by the phenomenal content, involving no subjective act. If this is true, then we cannot say that the integration of the subjective dimension is essential to the structure of conscious experience.

The problem with this objection is that it assumes, falsely in our view, that the subjective dimension of consciousness is something of which we are always explicitly self-aware. Given this assumption, it follows from the experience of absorbed coping that conscious experience does not always involve a subjective dimension. The assumption lacks phenomenological support, however. For we are in fact not always aware of the way in which we are conscious of the phenomenal content. We might even say that it is more the rule than the exception for the subjective effects to operate implicitly. A naive observer looking at a Necker cube, for example, will seldom notice that her subjective perspective on the cube correlates with the appearing cube until it is pointed out by the experimenter. The lack of explicit self-awareness, therefore, does not imply the absence of the subjective dimension in the experience.

For clarification, it may be helpful to describe the subjective dimension of conscious experience in general as *subjective attitudes*. Then we can draw a distinction between two forms of subjective attitudes, namely, between *subjective acts* of which the subject is explicitly self-aware while having the experience and *mere attitudes* which are invisible to the subject while the experience unfolds. On this terminology, the case of absorbed coping only shows that subjective acts are inessential for conscious experience; it does not show this about subjective attitudes in general. We therefore conclude that the case poses no threat to the claim that subjective attitudes and phenomenal contents are interdependent constituents of conscious experience.

If these phenomenological analyses are correct, we have good phenomenological reason to think that conscious experience is an integrated Gestalt that consists of a subjective attitude and a phenomenal content. When IIT addresses the integrated structure of consciousness, it only considers the integration internal to the phenomenal content. If the integration of the phenomenal content is best understood holistically, as

²³ However, Dow (2017) criticizes the general characterization of absorbed coping in terms of lack of self-awareness. This amounts to a different way of responding to the objection described in this paragraph.



we argued in the previous section, however, we should think of the integrated structure as extending into the subjective attitude as well.

5 A novel theoretical challenge for IIT

We have argued so far on phenomenological grounds that conscious experience is an integrated Gestalt constituted by a subjective attitude and a phenomenal content. But what difference does this make to IIT *qua* a scientific theory of the neural substrate of consciousness? We claim that our phenomenological critique poses IIT a novel theoretical challenge. To explain the challenge, however, we must introduce some ideas that figure in the other components of the theory, namely, the ontological postulates and the central identity.

It is in the ontological postulates that IIT develops its account of the neural substrate of consciousness. As we saw in the introduction, the main idea is that conscious experience is underpinned by a set of physical elements with the highest degree of differentiation and integration. More specifically, any physical system can be analyzed as a structure of complexes, that is, sets of physical elements with a degree of differentiation and integration or a Φ-value (Note that this is a different concept of "complex" from the mereological one discussed in section 2). Among them is a set of elements with the highest Φ -value (" Φ ^{max}"), which is called the *main complex* or the major complex of the system. Furthermore, each complex has its own conceptual structure—that is, each stands in a unique probabilistic relation to a different set of past and future states by virtue of the nature of its subcomponents and the interaction among them (e.g. Tononi 2015; Tononi et al. 2016). Different complexes with the same Φ-value will be differentiated by their conceptual structure. Thus the conceptual structure of the major complex varies as different sets of elements bear the highest Φ -value as the whole system changes its state through time. At any particular moment, however, it is the major complex of the system with its unique conceptual structure that underpins the corresponding conscious experience of that moment.²⁴

The other component of IIT, the central identity, concerns the relation between conscious experience and its physical substrate, or more specifically, the conceptual structure of the major complex. It claims that they are identical with one another. However, the sense in which they are identical remains ambiguous. Here is one of the recent formulations of the claim:

[T]he central identity of IIT can be formulated quite simply: an experience is identical to a conceptual structure that is maximally irreducible intrinsically [i.e., the conceptual structure of the major complex]. More precisely, a conceptual structure **completely specifies** both the quantity and the quality of experience: how much the system exists—the quantity or level of consciousness—is measured by its Φ^{max} value—the intrinsic irreducibility of the conceptual structure; which way it exists—the quality or content of consciousness— is specified by the shape of the conceptual structure. (Tononi and Koch 2015, p. 9, emphases in

²⁴ See Oizumi et al. (2014), pp. 3–15 for more detail.



italics in original, emphases in bold added; see also Tononi 2012, pp. 317-318, Oizumi et al. 2014, p. 3, Tononi 2015)

On the one hand, this allows a strong metaphysical reading, according to which conscious experience is nothing but the conceptual structure of the major complex. On this view, there is nothing more to being conscious than being a physical system with a particular Φ -value and a conceptual structure. This reading of the central identity is arguably the reason that some theorists including Tononi himself consider IIT to entail panpsychism.

One concern with this reading, however, is that there is almost nothing in the other components of IIT—the phenomenological axioms and the ontological postulates—that motivate this metaphysical claim. In the axioms and the postulates, IIT offers a general account of the physical mechanism of consciousness by focusing on its capacity to generate an integrated experience at the phenomenological level. What this implies about the metaphysical status of consciousness changes depending on the modality of the account. If the conceptual structure is considered to specify conscious experience by metaphysical necessity, then the account implies some form of physicalist panpsychism. If the specification relation is considered as nothing more than a contingent law of nature, the account implies some form of dualism. Either modal interpretation, however, is coherent with IIT's account of the mechanism of consciousness. This means that neither interpretation follows directly from the latter.²⁵

Although Tononi himself seems sympathetic to the panpsychic reading, accordingly, we will not consider this metaphysical commitment as an essential part of IIT. Instead, we adopt a more empirical reading of the central identity, according to which it is literally a claim about specification or systematic mapping. The question is whether there really holds a specification or mapping relation between the mathematical structure of the physical substrate and the phenomenological structure of conscious experience. Whether or not one agrees with the claim, this seems to be a sound empirical hypothesis inviting further scientific investigation.²⁶

Then we can think of the phenomenological analysis of the integrated structure of conscious experience developed so far as implying a novel theoretical challenge to IIT in the following way. The central identity taken as an empirical hypothesis consists of two claims: one about the possibility of specifying "the quantity or level of consciousness" from the Φ -value of the major complex; the other about the possibility of specifying "the quality or content of consciousness" from the conceptual structure. The former claim matters less for the current discussion.²⁷ The challenge from the

²⁷ See Bayne et al. (2016) for a critical assessment of the concept of the level of consciousness.



²⁵ As Peressini points out, the claim is only motivated by the observation that both the phenomenal content and the major complex bear an integrated, holistic structure of some sort, which he describes dismissively as "thin stuff on which to ground the [identity] claim" (Peressini 2013, p. 194).

²⁶ This is to say that we are going to read IIT as a theoretical framework that intends to unify phenomenological and neuroscientific accounts of consciousness, an answer to what Lutz and Thompson (2003) calls the "explanatory gap problem", as opposed to the hard problem of explaining the metaphysical relation between consciousness and physical nature (see Lutz and Thompson 2003, pp. 47–48). We are not alone in adopting such an empirical reading of IIT. For example, Tsuchiya et al. (2016) presumes an empirical interpretation of the central identity to suggest the possibility of testing the proposal by using the mathematical formalism of *category theory*. Similarly, Yoshimi (2011) suggests, independently of IIT, the possibility of identifying a "supervenience function" that links mathematically the dynamics of brain activity with that of conscious experience.

phenomenological critique emerges more in relation to the latter claim: If we are right that conscious experiences are integrated Gestalts consisting of a subjective dimension and a phenomenal content, then how should we consider the relation between the former and the physical substrate of consciousness? Given the central identity of IIT, it seems natural to posit a systematic mapping not only from the physical substrate onto the phenomenal content, but also onto the subjective dimension of the experience. The theoretical challenge is to explain how this might be the case.

There are a few dialectic options IIT can adopt in response. One option is to face up the challenge. IIT can update the current account of the physical substrate of consciousness so as to accommodate the subjective dimension into its explanatory scope. As the current version of the identity thesis holds that the phenomenal content maps onto the major complex, this update probably involves introducing a new theoretical construct that captures the aspect of the physical substrate that specifies the subjective dimension of conscious experience. Another option is to back up. One can drop the challenge, that is, by limiting the scope of the identity thesis to the phenomenal content of conscious experience.²⁸ Anthony Peressini, for example, argues that IIT is better understood as a theory only about the "qualitative character" (or the phenomenal content in our expression) of conscious experience, that is, as "integrated information theory of qualia (IITQ)" (Peressini 2013). Given the integrated structure at the phenomenological level, we are not sure if the neuroscientific account of the subjective dimension and that of the phenomenal content can really have different foundations. But this is an open question for further philosophical and empirical investigation. A third option is to dispute the phenomenology. One might deal with our challenge, that is, by denying that conscious experience is an integrated Gestalt constituted by a subjective dimension and a phenomenal content.

In any case, IIT does not afford to simply ignore the challenge. The most important implication from the phenomenological discussion is that whichever dialectic option one chooses to take, a serious consideration about the treatment of the subjective dimension is inevitable. If we are not going to back up or dispute the phenomenology, however, how could we possibly update IIT to accommodate the subjective dimension into its explanatory scope? The next section sketches one possible way to answer this question.

6 The noetic complex proposal

IIT holds that the major complex of the brain underpins conscious experience. The state of the brain at a particular moment, however, is not exhausted by the set of physical elements making up the major complex. This is made clear in one of the most recent presentations of IIT in terms of the distinction between *major* and *minor complexes* (Oizumi et al. 2014). While the former denotes the set of interacting elements with the locally highest Φ -value, the latter refers to other sets of interacting elements with some φ -value greater than zero. According to Tononi and his collaborators, "a system can condense [i.e., can be analyzed] into a major complex and minor complexes that may or may not interact with it" (ibid., p. 16).

²⁸ We thank the anonymous reviewer for reminding us of this option.



Given that minor complexes are also realizations of some capacity to integrate information, they are also considered as making some contribution to consciousness. In particular, it is suggested that minor complexes qualify as "minimally conscious" (ibid.). That is, the brain realizes more than one conscious experience by containing more than one complexes; however, we only become aware of the content realized by the major complex because it involves a larger quantity or level of consciousness compared to the other contents realized by the minor complexes. "In the healthy, adult human brain the qualia and Φ^{max} generated by the dominant main complex are likely to dwarf those specified by the minimally conscious minor complexes" (ibid.).

We find the introduction of the distinction between major complex and minor complex useful for better explaining the internal structure of the physical system underlying consciousness. However, we find it unilluminating to interpret minor complexes as specifying minimally conscious content of which we are completely unaware. It is an excessively speculative interpretation because no phenomenological evidence is available for the existence of minimally conscious contents in the first place; even worse, this is so precisely because of their definitive characterization as something beyond conscious awareness.

We therefore submit an alternative interpretation. We propose to develop the concept of minor complex so as to make it help IIT answer the theoretical challenge described in the last section—that is, the challenge to explain the neural substrate of the subjective dimension of consciousness. The proposal is to posit a specific class of minor complex that specifies the nature of the subjective dimension of the experience that is in place at the time. We shall call this class of minor complex specific to the subjective dimension *noetic complex*.

The proposal raises an immediate question. What are the differentiating features of noetic complexes? What makes them different from mere minor complexes that do not specify the subjective dimension of consciousness? We can start answering this question by focusing on the dynamic interplay between the subjective attitude and the phenomenal content at the phenomenological level described earlier (section 4.1).

Suppose IIT is right that the phenomenal content maps onto the major complex. If the subjective attitude maps onto the noetic complex, then, this gives some ground to suppose that the interaction between the subjective attitude and the phenomenal content at the phenomenological level maps onto the interaction between the noetic complex and the major complex at the system level. This is not so much an a priori philosophical analysis as an empirical conjecture open to empirical testing. Based on this conjecture, we can characterize the noetic complex, as a first approximation, as a minor complex that is particularly interactive with the major complex (Fig. 1).

We can be more specific by looking more closely at the interactions both at the phenomenological and the system level. At the phenomenological level, as we saw earlier, the subjective attitude affects the phenomenal content by affecting its self-organization, rather than instantly realizing a particular content. To see how such interaction between the two aspects of conscious experience can be specified at the system level, we only need to understand the interaction between the noetic complex and the major complex in terms of a switch between *attractors* from the perspective of dynamical systems theory. Attractors for a dynamical system are closed subsets of its phase space such that for many values of initial conditions, the system will evolve towards the subset (Guckenheimer and Holmes 1983; Lorenz 1963). Intuitively, if a



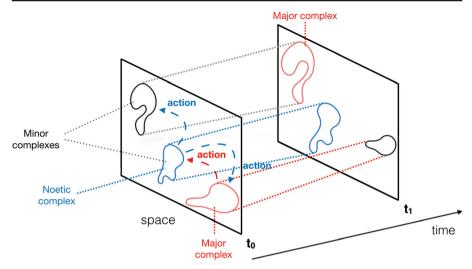


Fig. 1 Transition of a given physical system between states at time t_0 and t_1 . Complexes are represented as geometrical patterns, embedded in space and can be tracked across time. In this example, the major complex (red shape in the figure) has switched its location between t_0 and t_1 as a result of being perturbed by the noetic complex (blue arrow and blue shape in the figure). The noetic complex can also bear a different geometrical pattern by being perturbed by the major complex (red arrow in the figure). The noetic complex differs from mere minor complexes (black shape in the figure) by being highly interactive with the major complex

system evolves in time in such a way as to approach a particular state, then an attractor represents this state. An important feature of attractors for our purpose is that they need not be fixed. External perturbations to the system may push the whole system towards different attractors from its current one, but without making it instantly arrive in the new basin of attraction (Fig. 2). If this is how the noetic complex affects the spatial location of the major complex in the system, then there will be a systematic mapping between the interactions happening at the two levels. The mapping consists in how both the subjective attitude and the noetic complex affect their counterpart only by intervening on the latter's self-organization.

On the other hand, we also saw that the phenomenal content can affect the subjective attitude at the phenomenological level. To see how such interaction maps onto interactions at the system level, we only need to confirm that the noetic complex and the major complex are mutually influential—that is, that there truly is an *interaction*. If changes in the major complex affects the noetic complex by intervening on the latter's self-organization, just as the latter can affect the former in this way, then such interaction at the system level will specify at the phenomenological level modifications of the subjective attitude by the phenomenal content.²⁹

IIT hardly considers the possibility that parts of the system external to the major complex specify some aspect of conscious experience. The proposal that there is a special class of minor complex more interactive with the major complex than the rest,

²⁹ How does the distinction made at the phenomenological level between subjective acts and mere subjective attitudes map onto the structure of the physical system? One possibility is that the noetic complex intersects with the major complex when the subjective dimension figures in the phenomenal content at the phenomenological level, while the two complexes do not overlap when the subjective dimension remains outside the phenomenal content.



Lorenz Attractor

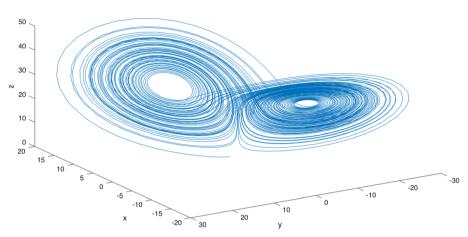


Fig. 2 Illustration of a switch between chaotic attractors (a.k.a. strange attractors; here the famous Lorenz attractor, first derived from a simple model of convection in the Earth's atmosphere (Lorenz 1963)). The blue line represents the evolution in time of the state of the system, and the attractors (or basins of attraction) lie at the center of the two concentric shapes. The system occasionally transfers from one basin of attraction to another without any abrupt jump, but by following a continuous path

however, is one that is already made in IIT (Oizumi et al. 2014, pp. 15–16). In this sense, it is not so much a radical divergence as a natural extension of IIT to introduce the concept of noetic complex. In fact, it is an empowering extension because in addition to identifying that conscious experience is underpinned by certain neural patterns, it can help IIT to further explain which aspects of the neural patterns underpin which aspects of the experience.

7 Conclusion

We examined the integrated structure of consciousness by developing a critique of IIT from the perspective of philosophical phenomenology. In our view, the integrated structure of the phenomenal content should be understood holistically in terms of integrated Gestalts. The argument for this can be summarized as follows:

- (a) If a whole is made up of interdependent constituents, it is an integrated Gestalt. [Sufficient condition for being an integrated Gestalt]
- (b) The phenomenal content is a whole made up of interdependent constituents. [Phenomenological analysis]
- (c) The phenomenal content is an integrated Gestalt. [From (a) and (b)]

Based on the holistic conception of integration, furthermore, we claimed that the integrated structure extends into the subjective dimension of consciousness. In other words, conscious experience is an integrated Gestalt that consists of a subjective attitude and a phenomenal content. The argument for this can be summarized as follows:



- (d) Conscious experience is a whole made up of a subjective attitude and a phenomenal content that are interdependent. [Phenomenological analysis]
- (e) Conscious experience is an integrated Gestalt made up of a subjective attitude and a phenomenal content. [From (a) and (d)]

This implies a novel theoretical challenge for IIT: It requires IIT to consider how to accommodate the subjective dimension into its explanatory scope. We sketched one possible way in which IIT might meet this challenge, calling it the noetic complex proposal. We believe this proposal offers a natural and empowering extension of the current version of IIT; however, we are also genuinely open to the possibility that there be other ways to answer the challenge. Whatever turns out to be the best response, most importantly, IIT cannot just ignore the challenge once and for all.

Acknowledgements The work of KM is supported by JSPS Overseas Research Fellowship (2016-2018), JSPS KAKENHIGrant Number 16K13147 and 18K00032. The work of OW is funded by the Earth-Life Science InstituteOrigins Network (EON) Research Fellow Program, supported by the John Templeton Foundation. KMthanks Hiroaki Hamada, Takuya Niikawa, and Satoshi Nishida for a close examination of the paper in itsearly stage of conception. Earlier versions of the paper have been presented at the Initiative for theSynthesis of the Studies of Awareness (ISSA) Summer School 2017, The Science of Consciousness2017, the Mind and Brain Annual Conference at the Center for Cognition and Neuroethics, and YHouseWorkshop "Opening Up Phenomenology". We appreciate the audience in each occasion for providing anumber of important feedbacks.

Author contributions The paper is a product of an ongoing conversation on the topic between the two authors. Sections 1-5 are single-authored by KM, and section 6 is jointly authored by KM and OW. Both authors edited and proof read the entire text, and they have a shared commitment to the content of the paper.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Arvidson, S. (2006). The sphere of attention: Context and margin. Dordrecht: Springer.

Barrett, A. B. (2014). An integration of integrated information theory with fundamental physics. *Frontiers in Psychology*, 5. https://doi.org/10.3389/fpsyg.2014.00063.

Bayne, T. (2010). The Unity of consciousness. Oxford: Oxford University Press.

Bayne, T., & Chalmers, D. J. (2003). What is the unity of consciousness? In A. Cleeremans (Ed.), *The Unity of consciousness: Binding, integration, and dissociation* (pp. 23–58). Oxford: Oxford University Press.

Bayne, T., Hohwy, J., & Owen, A. M. (2016). Are there levels of consciousness? *Trends in Cognitive Science*, 20(6), 405–413. https://doi.org/10.1016/j.tics.2016.03.009.

Beaton, M., & Aleksander, I. (2012). World-related integrated information: enactivist and phenomenal perspectives. *International Journal of Machine Consciousness*, 4(2), 439–455. https://doi.org/10.1142 /S1793843012400252.

Boly, M., Seth, A. K., Wile, M., Ingmundson, P., Baars, B., Laureys, S., Edelman, D. B., & Tsuchiya, N. (2013). Consciousness in humans and non-human animals: recent advances and future directions. Frontiers in Psychology, 4. https://doi.org/10.3389/fpsyg.2013.00625.

Carman, T. (2005). On the inescapability of phenomenology. In D. W. Smith & A. L. Thomasson (Eds.), *Phenomenology and philosophy of mind* (pp. 67–90). New York: Oxford University Press.

Cerullo, M. A. (2015). The problem with phi: a critique of integrated information theory. *PLoS Computational Biology*, 11(9), e1004286. https://doi.org/10.1371/journal.pcbi.1004286.

Chudnoff, E. (2013). Gurwitsch's phenomenal holism. *Phenomenology and the Cognitive Sciences*, 12(3), 559–578. https://doi.org/10.1007/s11097-012-9279-6.

Csikszentmihalyi, M. (1990). Flow: the psychology of optimal experience. New York: Harper & Row.



- de Barros, J. A., Montemayor, C., & Assis, L. P. (2016). Contextuality in the integrated information theory. In International Symposium on Quantum Interaction (pp. 57–70). Cham: Springer.
- Dow, J. (2017). Just doing what I do: on the awareness of fluent agency. *Phenomenology and the Cognitive Sciences*, 16(1), 155–177.
- Dreyfus, H. (2005). Overcoming the myth of the mental: how philosophers can profit from the phenomenology of everyday expertise. Proceedings and Addresses of the American Philosophical Association, 79(2), 47–65.
- Guckenheimer, J., & Holmes, P. (1983). Nonlinear oscillations, dynamical systems and bifurcations of vector fields. New York: Springer.
- Gurwitsch, A. (1964). The field of consciousness. Pittsburgh: Duquesne University Press.
- Gurwitsch, A. (1966a). Phenomenology of thematics and of the pure ego: Studies of the relation between Gestalt theory and phenomenology. Reprinted in *Studies in Phenomenology and Psychology* (pp. 175–286). Evanston: Northwestern University Press. (Original work published 1929).
- Gurwitsch, A. (1966b). A non-egological conception of consciousness. Reprinted in *Studies in phenomenology and psychology* (pp. 287–300). Evanston: Northwestern University Press. (Original work published 1941).
- Husserl, E. (1960). Cartesian meditations: An introduction to phenomenology. The Hague: Martinus Nijhoff (Original work published 1950).
- Husserl, E. (2012). Ideas: General introduction to pure phenomenology. New York: Routledge (Original work published 1913; original translation published 1931).
- Ihde, D. (2012). Experimental phenomenology: Multistabilities. Albany: SUNY Press.
- Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: two distinct brain processes. Trends in Cognitive Science, 11(1), 16–22.
- Kriegel, U. (2015). The varieties of consciousness. New York: Oxford University Press.
- Lorenz, E. N. (1963). Deterministic non-periodic flow. Journal of the Atmospheric Sciences, 20(2), 130-141.
- Lutz, A., & Thompson, E. (2003). Neurophenomenology: integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies*, 10(9–10), 31–52.
- Merleau-Ponty, M. (2012). *Phenomenology of perception*. New York: Routledge (Original work published 1945).
- Montemayor, C., & Haladjian, H. H. (2015). *Consciousness, attention, and conscious attention*. Cambridge: The MIT Press.
- Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *PLoS Computational Biology*, 10(5), e1003588. https://doi.org/10.1371/journal.pcbi.1003588.
- Peressini, A. (2013). Consciousness as integrated information: a provisional philosophical critique. *Journal of Consciousness Studies*, 20(1–2), 180–206.
- Prinz, J. (2012). The conscious brain: How attention engenders experience. New York: Oxford University Press.
- Schwitzgebel, E. (2015). If materialism is true, the United States is probably conscious. *Philosophical Studies*, 172(7), 1697–1721. https://doi.org/10.1007/s11098-014-0387-8.
- Tegmark, M. (2016). Improved measures of integrated information. *PLoS Computational Biology, 12*(11), e1005123. https://doi.org/10.1371/journal.pcbi.1005123.
- Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, 5(1), 42. https://doi.org/10.1186/1471-2202-5-42.
- Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *The Biological Bulletin,* 215(3), 216–242. https://doi.org/10.2307/25470707.
- Tononi, G. (2012). Integrated information theory of consciousness: an updated account. *Archives Italiennes de Biologie, 150,* 290–326.
- Tononi, G. (2015). Integrated information theory. Scholarpedia, 10(1). https://doi.org/10.4249/scholarpedia.4164.
- Tononi, G., & Koch, C. (2015). Consciousness: here, there and everywhere? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*, 20140167. https://doi.org/10.1098/rstb.2014.0167.
- Tononi, G., Boly, M., Massimini, M., & Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. *Nature Reviews Neuroscience*, 17, 450–461. https://doi.org/10.1038/nrn.2016.44.
- Tsuchiya, N., Taguchi, S., & Saigo, H. (2016). Using category theory to assess the relationship between consciousness and integrated information theory. *Neuroscience Research*, 107(1–7), 1–7. https://doi. org/10.1016/j.neures.2015.12.007.
- Varzi, A. (2016). Mereology. In Edward N. Zalta (Ed.), The Stanford encyclopedia of philosophy (Winter 2016 edition). https://plato.stanford.edu/archives/win2016/entries/mereology/. Accessed 28 November 2018



Wiese, W. (2017). How to solve the problem of phenomenal unity: finding alternatives to the single state conception. *Phenomenology and the Cognitive Sciences*, 16(5), 811–836. https://doi.org/10.1007/s11097-016-9478-7.

Wrathall, M. A. (2014). Introduction: Hubert Dreyfus and the phenomenology of human intelligence. In H. L. Dreyfus (Ed.), *Skillful coping: Essays on the phenomenology of everyday perception and action* (pp. 1–22). New York: Oxford University Press.

Wu, W. (2014). Attention. New York: Routledge.

Yoshimi, J. (2011). Phenomenology and connectionism. Frontiers in Psychology, 2. https://doi.org/10.3389/fpsyg.2011.00288.

Affiliations

Katsunori Miyahara 1,2 · Olaf Witkowski 3,4

- School of Humanities and Social Inquiry Bldg. 19, University of Wollongong, Northfields Avenue, Wollongong, NSW 2522, Australia
- Graduate School of Arts and Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo, NSW 153-8902, Japan
- Earth-Life Science Institute (ELSI), Tokyo Institute of Technology, 2-12-1-IE-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan
- Institute for Advanced Study (IAS), 1 Einstein Drive, Princeton, NJ 08540, USA

