Minimal phenomenal experience

Meditation, tonic alertness, and the phenomenology of “pure” consciousness

Thomas Metzinger* (metzinger@uni-mainz.de)

Abstract

This is the first in a series of instalments aiming at a minimal model explanation for conscious experience, taking the phenomenal character of “pure consciousness” or “pure awareness” in meditation as its entry point. It develops the concept of “minimal phenomenal experience” (MPE) as a candidate for the simplest form of consciousness, substantiating it by extracting six semantic constraints from the existing literature and using sixteen phenomenological case-studies to incrementally flesh out the new working concept. One empirical hypothesis is that the phenomenological prototype of “pure awareness”, to which all such reports refer, really is the content of a predictive model, namely, a Bayesian representation of tonic alertness. On a more abstract conceptual level, it can be described as a model of an unpartitioned epistemic space.

Keywords

Consciousness as such ∙ Dharmakāya ∙ Empty cognizance ∙ Minimal model explanation ∙ Minimal phenomenal experience ∙ Minimal phenomenal selfhood ∙ Pure awareness ∙ Pure consciousness ∙ Rigpa ∙ Sākṣin ∙ Samadhi ∙ Self-luminosity ∙ Tonic alertness ∙ Transparency ∙ Turiya ∙ Wakefulness ∙ Ye shes

This article is part of a special issue on “Radical disruptions of self-consciousness”, edited by Thomas Metzinger and Raphaël Millière.

This paper is the first in a series of publications trying to develop a new approach to the problem of consciousness. In section 1 I will give a short introduction to the general idea, the strategy of explanation by minimal models, and list some of my epistemic goals. Section 2 looks at the phenomenology of “pure consciousness” by using a selection of case studies and developing a set of six semantic constraints for a finer-grained concept of “minimal phenomenal experience”. Section 3 provides

*Johannes Gutenberg University, Mainz
some conceptual differentiation, distinguishing between the phenomenal character of “wakefulness” and the functional property of tonic alertness. The last section presents a new hypothesis about the experience of consciousness as such: It is a subsymbolic, Bayesian, and aperspectival representation of tonic alertness.

1 Introduction

1.1 From MPS to MPE

In 2009, Olaf Blanke and Thomas Metzinger asked two questions: What are the minimally sufficient conditions for the appearance of a phenomenal self, that is, the fundamental conscious experience of being someone? What are necessary conditions for self-consciousness in any type of system (Blanke & Metzinger, 2009, p. 7)? One of their findings was that bodily agency is a causally enabling but not a constitutive condition for phenomenal selfhood (Blanke & Metzinger, 2009, p. 13). Blanke and Metzinger then introduced the concept of “minimal phenomenal selfhood” (MPS), the central defining features of which are (i) a globalized form of identification with the body as a whole (as opposed to mere ownership for body parts), (ii) spatiotemporal self-location, and (iii) a weak, geometrical first-person perspective (1PP).

This paper begins to develop a related strategy for “minimal phenomenal experience” (MPE; this term was originally introduced in Windt, 2015b). It will take as a new entry point not full-body illusions, but the specific phenomenology of “pure consciousness” in meditation. On the phenomenological level of analysis, part of the working hypothesis is that MPE lacks (i) MPS and all other forms of egoic self-consciousness, (ii) time representation, and (iii) a spatial frame of reference. Where for MPS the global phenomenology of identifying with the body as a whole was the empirical starting point, here, I begin with all reports in which subjects claim that they have actually had an experience of consciousness as such. Can a conscious system be exclusively aware of awareness itself? If yes, would there be a specific form of phenomenal character that is instantiated during such episodes? And if there really exists something like the simplest form of conscious experience – what are the minimally sufficient conditions for MPE to occur in neurotypical humans, and what would be necessary conditions for the experience of consciousness as such to appear in any type of system?1

1As a starting point, I have here mostly chosen reports about “pure awareness” or “pure consciousness” from contemplative practitioners, but in principle there may be many other relevant kinds of minimal phenomenal experience, for example as caused by subanesthetic doses of general anesthetics (Russell, 2013; Sanders, Tononi, Laureys, & Sleigh, 2012) or as experienced by patients moving in and out of coma (Giacino et al., 2002; Laureys, Owen, & Schiff, 2004). My reason for selecting this highly specific subset of contemplative experiences is twofold. First, if we are interested in arriving at a minimal-model explanation for phenomenal consciousness, it is quite simply the most natural first step to closely investigate all those states in which subjects actually report experiencing the “essence” of consciousness or consciousness as such. Second, what makes the...
1.2 Scientific explanation by minimal models

One central motivation behind the current initiative is to test a fresh alternative to the now-classical research strategy of isolating the “neural correlate of consciousness” (NCC, Metzinger, 2000). The new idea is to develop a “minimal model” of conscious experience. In the present context, developing a “minimal model” of conscious experience is something very benign and modest. For example, given the current state of consciousness science, a minimal model will not yet involve elaborate formal procedures like mathematically delimiting a universality class (Batterman & Rice, 2014). Rather, it can be seen as a heuristic strategy of minimalist idealization:

- Constructing and investigating a model of conscious experience that includes only the core causal factors giving rise to the target phenomenon;
- Including only causal factors that make a difference to the actual occurrence and the essential phenomenal character itself;

The phenomenology of certain meditative states interesting from a scientific perspective is that the philosophical phenomenologies that typically come along with them claim a conjunction of minimal dimensionality (i.e., an “empty” form of conscious experience lacking internal structure and content) and maximal prototypicality (i.e., significant if often-unnoticed overlap with most other exemplars of the category of “consciousness”).

I presuppose that the concept of phenomenal consciousness has not definitional but probabilistic structure. “Probabilistic structure” means that a certain conscious experience falls under MPE if it satisfies a sufficient number of constraints encoded by MPE’s semantic constituents (Margolis & Laurence, 2014). In principle, there could exist multiple and loosely overlapping sets of sufficient conditions. Ultimately, MPE is a cluster concept and “MPE” refers to a phenomenological prototype. This implies that there is no set of necessary and sufficient conditions that allows us to say whether a representational state is conscious or not. Rather, “consciousness” refers to states, regions, or trajectories in a multidimensional space constituted by functional and content-related dimensions – it is a heterogeneous and graded construct, and membership comes in degrees (Bayne, Hohwy, & Owen, 2016; for a similar approach to the concept of “mind wandering”, see Seli, Kane, Metzinger, et al., 2018). The underlying dimensions which give a structure to the space of conscious experience create a complex pattern of family resemblances, and absolute orderings may not be possible in all cases. There will be degrees of prototypicality, because highly prototypical members have attributes that overlap with most other exemplars of the category, while low-prototypical members have little overlap. Membership in the family of phenomenal states will therefore be graded, and quantitative assessments of prototypicality may actually help us understand why some phenomenological exemplars capture the intuitive “essence” of consciousness better than others.

Please note how this implies a radical departure from essentialist intuitions as might seem to be automatically implied by the two questions about necessary and sufficient questions in section 1.1: It may not be possible to isolate the simplest form of conscious experience in terms of one unique minimum of dimensionality for the relevant state-space (or an individual state displaying the minimal set of characteristic features), and for a given system multiple low-dimensional states may actually be possible or even co-exist. For example, in a 1,000,000-dimensional state space there might be five different kinds of three-dimensional states, each of them exhibiting a unique phenomenal character.
- developing an idealized model of **universal and repeatable features** serving to gradually isolate the fundamental, explanatorily relevant, and structurally stable properties that underlie different forms of conscious experience.

The opposite of a minimal model would be a *maximal* model (i.e., a model which would be “ideal” in a very different sense): a theory which would capture all the behaviour, both repeatable and unrepeatable, of a given system. For complex systems like the conscious human brain, this traditional modelling strategy is an unattainable goal; many of the fine details will actually get in the way of scientific explanation and detract from a deeper understanding of the research target (Batterman, 2002, p. 22). However, if the phenomenon to be explained is a large-scale regularity that is largely independent of micro-scale details and features of individual elements, it can prove helpful to identify the *minimal conditions* necessary for the explanatory target to occur, later linking them back to the original phenomenon of interest (for three historically important case studies in ancestral heredity, the ideal gas law in physics, and evolutionary biology, see section 2 in Rice, Rohwer, & Ariew, 2019; see also footnote 4).

In the present context, there are two major aspects of the minimal-model strategy. The first is to describe the target phenomenon in an uncluttered way, abstracting away from everything that is not an essential feature of the core explanandum. If the systems of interest are conscious systems, three questions arise: (i) whether *phenomenality as such* has a distinct experiential character, (ii) whether this character ever occurs in isolation, and (iii) whether it can be described in a conceptually precise manner (this is the main goal of the current paper). The second aspect aims at an economical strategy for describing the network of causal factors giving rise to this specific explanandum – given that a precise description succeeds. Here, constructing a *minimal* model would amount to eliminating superfluous details by extracting only the explanatorily relevant causal structure underlying the experience of *phenomenality as such*, thereby providing a medium-level functional analysis while still omitting most of the structural details of a fully mechanistic explanation (i.e., a “mechanism sketch”; cf. Piccinini & Craver, 2011). Both aspects could come together if mathematically stable features of the physical dynamics characterizing the global neural correlate of MPE in humans could be directly mapped to corresponding abstract patterns within phenomenal dynamics independently of potentially unknowable and fine-grained neuroscientific details, thereby yielding what we are most interested in: universal and reliably *repeatable* properties instantiated by all forms of subjective experience.

The first positive proposal of this paper is that individual episodes of MPE could be instantiations of such an abstract pattern, perhaps even of a universal or at least strongly prototypical feature. In terms of research heuristics, my second proposal is that it is simply natural to focus first on exactly those states (long identified by contemplative traditions) which are described as episodes of “pure awareness”, “empty cognizance”, or “consciousness as such”. *Prima facie* and given the vastness...
of our own phenomenal state space, this may be the most direct route to studying such an abstract pattern without any distracting additions.

The strategy of “minimalist idealization” looks for the decisive explanatory and computational factors by first attempting to construct exactly solvable models (Batterman, 2002, p. 37). Therefore, minimal models do not aim at maximal fidelity or completeness; they are not yet mappings to fine-grained functional mechanisms because the accuracy conditions for an actual representation of causal pathways are not intended to be very strong (Weisberg, 2012, Chapter 6, section 6.1.2). However, explanation by minimal models may nevertheless involve a phase of de-idealization at a later stage, turning them into a continuing research program eventually leading to a multilevel mechanistic explanation (Piccinini & Craver, 2011). This is one of my longer-term goals. But I also believe that in order to really solve the problem of consciousness we will ultimately need a theory of MPE – because only a minimal model can give us a deep scientific understanding of the essence of phenomenal experience.

1.3 Sketching the epistemic goal of the MPE approach

There are five factors that make the concept of “MPE” interesting. First, MPE could refer to the common phenomenological denominator that is always present whenever there is conscious experience at all, however rich or limited. Second, a concerted attempt to isolate this hypothetical common denominator could lead to interdisciplinary unification in consciousness research by connecting different domains and experimental approaches in a new way. Third, by aiming at a minimal-model explanation we could arrive at an entirely new theoretical model of conscious experience itself, providing us with a fresh perspective and possibly even a new window onto the target. Fourth, the MPE approach could lead to unification and considerable enrichment within analytical philosophy of mind itself, by elevating comparative and transcultural philosophy of mind to a new and more systematic level. And finally, reports about “pure consciousness” experiences are a neglected empirical phenomenon worth studying in its own right.

From now on, I will use the phenomenological notions of “pure awareness” and “pure consciousness” interchangeably. “MPE” refers to the theoretical concept of minimal phenomenal experience as it could figure in a minimal-model explanation of consciousness. I will here limit myself to one plausibly homogeneous subclass of MPE: to a paradigmatic form of “pure consciousness” experience as it appears in some forms of contemplative practice (see footnote 1).

If one accepts the relevance of the five reasons given above, then the most pressing epistemic goal consists in first getting the phenomenology right. The guiding phenomenological question is this: Does pure awareness have a minimal, distinct, and unique phenomenal character sui generis?
We can also ask four types of slightly more detailed question:

(1) Is there a form of consciousness that does not instantiate minimal phenomenal selfhood (Blanke & Metzinger, 2009) or an epistemic agent model (Metzinger, 2017, sec. 2.5)?
   
   → Are selfhood (i.e., egoic self-consciousness, the experience of “being someone”) and perspectivalness (i.e., an individual and consciously experienced first-person perspective) necessary conditions for phenomenality?
   
   → If they are not, what is the epistemological status of reports about such episodes?

(2) Does “pure” consciousness really exist?
   
   → Are there any concrete examples of phenomenality without intentional content?
   
   → Are there types of conscious experience which do have content, but of which we would predict that they will be reported as “empty” or “contentless” forms of experience?

(3) Can we develop a minimal model for conscious experience as such?
   
   → Is there a single kind of conscious experience that can count as the simplest instance of the target phenomenon?
   
   → Fundamentality: Is there a fundamental (and perhaps implicit) kind of phenomenal character sui generis, which can at times be made explicit and which underlies or “permeates” all other forms of phenomenal experience?

(4) Is there one maximally abstract phenomenal property?
   
   → Is “pure awareness” a unique form of phenomenal character of which we could say that it forms a phenomenological kind of its own and that it subsumes all other forms?

I will present preliminary answers to some of these questions in section 4. First, let us penetrate a little deeper into the phenomenological landscape surrounding MPE, using a set of fourteen case studies plus two canonical phenomenological descriptions from the relevant literature.

2 MPE on the phenomenological level of analysis

One major research target for the current initiative is the development of a fully validated psychometric instrument for MPE. In this introductory paper I will take three steps. As a first, preparatory step for creating a finer-grained, more evidence-based description I will now treat MPE as an exclusively phenomenological notion, deliberately leaving aside all the metaphysical interpretations and theoretical background assumptions frequently offered along with experiential reports by
humankind’s contemplative traditions. Second, I will attempt to semantically enrich the new concept by extracting six general constraints from both Eastern and Western philosophical traditions, sometimes mixing phenomenological analysis and exegesis. Accordingly, we will arrive at a preliminary 6D-model of what has hitherto been called “pure” consciousness. The third step will have to consist in further increasing the dimensionality of our conceptual instruments, by asking actual meditators whether they have ever had the relevant kind of “pure awareness” experience and if so, how they would describe its phenomenal character (cf. footnote 1).

The six constraints to be introduced in the following subsections are:

- **Wakefulness (PC1)**
- **Low Complexity (PC2)**
- **Self-Luminosity (PC3)**
- **Introspective Availability (PC4)**
- **Epistemicity (PC5)**
- **Transparency/Opacity (PC6)**

### 2.1 Pure consciousness in contemplative phenomenology

MPE is probably an experience every human being knows. Nonetheless, most explicit reports come from contemplative traditions, and I will therefore treat it for the most part as a specific and very simple subset of meditative experiences. In his classic monograph *Philosophy and Mysticism* Walter Terence Stace categorizes reports of “pure consciousness” as a specific, namely introvertive, type of mystical experience:

> There would be no mental content whatever but rather a complete emptiness, vacuum, void. One would suppose *a priori* that consciousness would then entirely lapse and one would fall asleep or become unconscious. But the introvertive mystics – thousands of them all over the world—unanimously assert that they have attained to this complete vacuum of particular mental contents, but that what then

---

2It must be clearly noted that the MPE approach as sketched out here operates under naturalistic background assumptions, and it does so without having given an independent argument to support these assumptions. Second, in some relevant philosophical traditions there have been very different conceptual, epistemological, metaphysical, and soteriological interpretations of the relevant phenomenological reports than the one presented here, in particular by those scholar-practitioners of the past who extensively cultivated and subsequently underwent such experiences themselves (see, e.g., Higgins, 2013, p. 10, footnote 4). The philosophical difficulty lies in correctly assessing the phenomenology of certainty and insight *with respect to the states’ own metaphysical status* that often goes along with such states, especially when they involve permanent alterations to conscious experience (see section 3.1 in Metzinger & Windt, 2015; Metzinger, 2014). On this second point, an anonymous reviewer remarked that for at least some types of MPE experiences there is an important aspect of the phenomenological profile that should not be ignored without argument, namely their “noetic purport to be metaphysically ultimate”. For recent alternatives to naturalism, see for example Albahari (2020); Finnigan (2017). See also footnote 15.
happens is quite different from a lapse into unconsciousness. On the contrary, what emerges is a state of pure consciousness – “pure” in the sense that it is not the consciousness of any empirical content. It has no content except itself. [...] It is the bare unity of the manifold of consciousness from which the manifold itself has been obliterated. This seems analogous to saying that if from a whole or unity of many parts we could subtract all the parts, the empty whole or unity would be left. (Stace, 1960, p. 86)

Similarly, in a well-known collection of texts, first published by Oxford University Press in 1990 under the title The Problem of Pure Consciousness, editor Robert Forman defines “Pure Consciousness Events” (PCEs) as a “wakeful though contentless (non-intentional) consciousness”:

One neither thinks nor perceives any mental or sensory content. Yet, despite this suspension of content, one emerges from such events confident that one had remained awake inside, fully conscious. This experience, which has been called the pure consciousness event, or PCE, has been identified in virtually every tradition. [...] The pure consciousness event may be defined as a wakeful but contentless (non-intentional) consciousness. (Forman, 1998, pp. 185–186)

3Brentyn Ramm has recently introduced the closely related concept of a “Pure Awareness Experience” or “Pure Awareness Event” (PAE), which allows him to distinguish between “Objectless Pure Awareness Experience” and “Object-Directed Pure Awareness Experience”, cf. Ramm (2019, p. 9). Note how this then allows for two further distinctions, which are relevant and already well known in the contemplative literature: “Subject-Directed Pure Awareness Experience” and “Self-Directed Pure Awareness Experience”, for examples see Harding (2000, p. 9); Kelly (2019, p. 100).

The last phenomenological notion is directly related to the philosophical project of formulating a conceptually clear first-order reflexive model of MPE, cf., e.g., Montague (2016, p. 63). Ramm’s distinction itself points to the gradualist approach I pursue by introducing the LOW-COMPLEXITY constraint below, which will hopefully allow for finer-grained phenomenological distinctions in the future.

4Interestingly, the same author also endorses something strongly resembling the “minimal-model approach” introduced in section 1.2 above:

When a biologist seeks to understand a complex phenomenon, one key strategy is to look at it in its simplest form. Probably the most famous is the humble bacterium E. coli. Its simple gene structure has allowed us to understand much of the gene functioning of complex species. Similarly, many biologists have turned to the “memory” of the simple sea slug to understand our own more kaleidoscopic memory. Freud and Durkheim both used totemism, which they construed as the simplest form of religion, to understand the complexities of religious life. The methodological principle is: to understand something complex turn to its simple forms. Mystical experiences may represent just such a simple form of human consciousness. [...] These PCEs, encounters with consciousness devoid of intentional content, may be just the least complex encounter with awareness per se that we students of consciousness seek. The PCE may serve, in short, as the E. coli of consciousness studies. (Forman, 1998, pp. 185–186)
From these classical sources, we can already derive the first two major phenomenological constraints (PCs):

- **Wakefulness (PC1):** the phenomenal character of tonic alertness (see section 3.1).

- **Low Complexity (PC2):** often described as the complete absence of intentional content, in particular of high-level symbolic mental content (i.e., discursive, conceptual, or propositional thought), but also of sensorimotor or affective content.

If we accept these first two constraints, then a theory of MPE will have to be a theory of contentless, restful alertness – a form of phenomenal experience that in its purest form does not instantiate any intentional properties. From a philosophical perspective, this conjunction of two necessary conditions is simple and straightforward. As a first characterization, it is certainly helpful.

However, please note how even the very first classical quotations presented above introduce a whole range of conceptual ambiguities. In claiming that “[i]t has no content except itself”, Walter Stace (1960, p. 86) introduces reflexivity as an additional constraint for a conceptual prototype of pure consciousness; when he speaks of an “empty whole or unity”, one is confronted with the problem that on closer inspection the phenomenology of wholeness and the phenomenology of unity are very different (Wiese, 2018). And if Robert Forman defines PCEs by making use of the notion of “(non-intentional) consciousness”, this immediately creates the need for a background theory about what “intentional” content really is, how many kinds of mental content there are, and how phenomenal and intentional properties are related. In particular, these conceptual ambiguities leave one specific possibility open: The absence of all empirical content could itself be an appearance, and what subjects (falsely) describe as a “contentless” phenomenal state could actually carry an abstract form of intentional content. If there is a content that can account for the reported experience of contentlessness, then Representationalism remains tenable. This is the line I will be pursuing.

From a very different perspective, here is how Jonathan Shear approaches the point:

> Of course, as empirically plausible models of the human brain provide us with clear examples of non-conceptual abstraction, “abstractness” does not imply conceptual or propositional content. In addition, whenever phenomenological reports are the target of analysis, such reports can be doubted. As Uriah Kriegel puts the point:

> Such a state of pure, contentless awareness is the state that certain meditative techniques are supposed to lead to. Certainly the state in question has been described along these lines by some practitioners. It is nonetheless open to the theoretician to challenge the practitioner’s own description of the state she is in. (Kriegel, 2009, p. 175, footnote 32)
Experience of the deepest level, that of pure consciousness awake to its own nature in and by itself, is generally held to be especially important. The defining characteristic of this experience is the complete absence of all sounds, tastes, thoughts, feelings, images, and anything else that one can ever imagine. [...] But they have in common the idea that it is possible for all empirical content to disappear, while one nevertheless remains awake.

What then is the experience like? By all accounts it is not like anything. One can have it and remember it – one knows that one was not asleep. But one does not remember it as anything at all. It is just itself – unimaginable and indescribable. The experience itself is extraordinarily abstract. Indeed, it is the logical ultimate of abstraction, since by all accounts it is what remains after everything that can possibly be removed from experience has been removed, while one nevertheless remains awake.

It should be noted, however, that even if the natural response to this experience is to describe it as “contentless”, it is still appropriate to raise the question of whether it is in fact completely contentless. For while experiences [...] might be so subtle and abstract that they naturally seem to be completely contentless, they might nevertheless actually have some, albeit very abstract, content. (Shear, 2007, pp. 700, 702–703)

If we are searching for a positive characterization of the target, an additional difficulty is that Low Complexity already introduces one exclusively negative element. And indeed, if one looks at the relevant literature, it is striking how many phenomenological characterizations of MPE episodes are of an exclusively negative kind. PCEs are described as:

- **Non-sensory**: MPE itself instantiates no perceptual qualities. This also includes “no-thingness”, i.e., the absence of the phenomenal property of “objecthood”. This means that there is no subjective experience of distinct multimodal objects as integrated from different sensory features and as segmented from a background or perceptual scene, being juxtaposed or “thrown against” a perceiving self;

- **non-motor**: absolute stillness, no motion in space;

- **atemporal**: an absence of temporal experience, no motion in time;

- **non-cognitive**: non-symbolic and non-conceptual, no mind-wandering;

- **non-egoic**: no self-location in time, no self-location in space, no quality of agency (either mental or bodily), no localized unit of identification (UI6), i.e.,

---

6The “unit of identification” (UI) is whatever form of experiential content leads to phenomenologi-
MPE itself is not even characterized by MPS, the minimal phenomenal sense of selfhood;

- **unbounded**: there is no second, finite region to which attention could be directed, and there are no consciously experienced boundaries, limits, or a horizon;

- **aperspectival**: no epistemic agent model (or EAM\(^7\)), no passive personal-level self-as-subject.

Hence, phenomenologically, the relevant phenomenal character seems not to be a **first-person** character: during its actual occurrence it is not tied to an individual first-person perspective (this feature is often called “non-duality” in Buddhist or Neo-Advaitan literature).

_Wakefulness_ is by far the most frequently mentioned kind of phenomenal character in descriptions of “pure consciousness”, and this observation is an important premise for the argument to be developed later. The second most frequent element is an absence of spatiotemporal self-location. “Spatial self-location” means that one experiences oneself as located at a point or a region in a spatial frame of reference (typically, as embodied), while “temporal self-location” means that one is _present_ in a temporal order, in a consciously experienced “now” (Blanke & Metzinger, 2009). “Nowness” itself is a form of temporal experience, and this already leads to two conclusions that can guide future neuroscientific research into the underlying correlates: Subjects experiencing the relevant states are not _oriented_ in time and space, and the processing of temporal information has been transiently suspended. To demonstrate the first and second aspect, here are three case studies:

**Case study #1.** I experience pure consciousness as a state of unbound- edness and total ease and deep relaxation. There are no thoughts, no feelings, or any other sensations like weight or temperature. I just know I am. There is no notion of time or space, but my mind is fully awake and perfectly clear. It is a very simple and natural state. (Alexander, 1988, p. 3)

\(^7\)An EAM is a special layer in the human self-model. It is a conscious internal model of the self as actively selecting targets of knowledge, as a mental agent who stands in epistemic relations (like “perceiving”, “believing”, “knowing”) to the world and to itself (as in “controlling the focus of attention”, “reasoning about the world and oneself”, or “knowing that one knows”), and as an entity that has the capacity to actively _create_ such relations of knowing. The EAM is the origin of our consciously experienced first-person perspective. For more see Metzinger (2017, sec. 2.5).
Case study #2. [...] a state of complete rest, full consciousness without content and unbounded in time and space. (Severeide, 1990, p. 1570; quoted from Bachmann, 2014, p. 52)

Case study #3. [...] a couple of times per week I experience deep, unbounded silence, during which I am completely aware and awake, but no thoughts are present. There is no awareness of where I am, or the passage of time. I feel completely whole and at peace. (Travis & Pearson, 2000, p. 81)

The third negative phenomenological descriptor which I want to highlight is “unboundedness”. It is present in all three case studies presented above, and it also frequently occurs in the literature. We must avoid a phenomenological non-sequitur at this point: To say that the space of pure consciousness is “boundless” does not imply that there is an explicit experience of infinite expansion or of large distances. Rather, it means that there is no “other side beyond the boundary” to which attention could shift. To give an example, in visual awareness attention can shift from a red colour patch into an adjacent green colour patch, transgressing a chromatic boundary. For the phenomenology of tonic alertness there is no such boundary, because everything outside the space of wakefulness cannot be deliberately attended to. It is simply unconscious.

Prima facie this has the interesting consequence that the method of phenomenal contrast (Siegel, 2007, p. 135) cannot be applied to the phenomenology of tonic alertness. It seems impossible to point to a pair of two experiences which differ only in terms of the relevant content; if there is no phenomenology of wakefulness, then there simply is no conscious experience altogether. In other words, if prototypical MPE as characterized above really is consciousness as such, then it makes no sense to contrast its specific phenomenal character with what could only be unconscious states devoid of such character. However, because for most of us the phenomenology of tonic alertness is normally fully transparent, this point will always remain strongly counterintuitive.

In any case it would be false to interpret relevant reports of “unboundedness” as the phenomenology of an explicitly experienced and merely quantitative infinity, or as a concrete, endless expanse of some sort. It just means that a) the phenomenal character in question includes the potential for expansion, and b) that in the experience of pure awareness itself there is no such thing as a “beyond” – another consciously experienced finite region or realm “on the other side” of a boundary. This has a second implication. In the relevant phenomenological domain, there are no countable entities: The domain is an unstructured space which has neither centre nor periphery and which cannot be introspectively individuated to create indivisible units of experience – not even a single one. For our growing list of phenomenological descriptors for MPE, this has the interesting consequence of

If this conclusion is valid, then describing inner awareness as peripheral ultimately rests on a misguided spatial metaphor, cf. Kriegel (2009, p. 176).

© The author(s). https://philosophymindscience.org ISSN: 2699-0369

The first theoretical conclusion was that the method of phenomenal contrast may not be applicable to MPE. The second conclusion now is that the relevant phenomenal character refers to the conscious experience of an unstructured space without a centre or boundaries, which also lacks the phenomenal quality of singularity or numerical identity. Against this background of predominantly negative characterizations there is a third conclusion that we can draw about Wakefulness and Low Complexity: There will be variable degrees of constraint satisfaction. The concept of “MPE” really refers to a multidimensional space of possible conscious states, a space which may contain different points, regions, or trajectories (Bayne et al., 2016). Prima facie, judging from the currently available set of case studies, there clearly seems to be a primary dimension constituted by the phenomenal character we refer to by the folk-psychological term “wakefulness”, and it is this phenomenological dimension that should become the first target of empirical studies. “Wakefulness as such” seems semantically related to similar phenomenological notions like “mental clarity” (but without mental content), “cognitive lucidity” (but without conscious cognition), or “bare awareness” (without object). What all of these have in common is that they refer to a kind of conscious experience that can clearly come in different strengths and degrees of intensity. “MPE” refers to their common phenomenal denominator.

Let me give two examples. First, it is plausible to assume different degrees of wakefulness phenomenologically characterizing such states. Subjectively, there should be an introspectively detectable gradient of tonic alertness. For example, satisfaction of Wakefulness could be higher during certain meditative or drug-induced states, but lower during episodes of mental fatigue or sleep deprivation, or when we are close to sleep onset or fainting.

To make this even clearer, let me briefly introduce a conceptual distinction which will be explained in detail in section 3.1. From now on, “wakefulness” exclusively refers to the phenomenal experience of tonic alertness, and not to the functional property of alertness as such or to the concrete physical property of cortical arousal. It is important never to conflate “wakefulness” as referring to a phenomenal property (that is, the conscious experience of sustained, cue-independent, tonic alertness as such) with “wakefulness” as a functional property (for example, as physically realized by the level of cortical arousal). Not only is it logically possible to have functional wakefulness without conscious experience (i.e., “wakeful unawareness”), but the possibility is also empirically demonstrated by the existence of a wake/sleep cycle in some types of coma patients, for example in the minimally conscious state (MCS) or in unresponsive wakefulness syndrome (UWS), tellingly termed Wachkoma in German (Laureys et al., 2010). In addition, it is empirically plausible that the conscious experience of wakefulness can exist in low-arousal states like NREM sleep, as shown by the existence of lucid dreamless sleep (Dentico et al., 2016; Ferrarelli et al., 2013; Maruthai et al., 2016; Mason et al., 1997;
Mason & Orme-Johnson, 2010; Metzinger, 2019, sec. 5; Thompson, 2015a, 2015b; Windt, 2015a; Windt, Nielsen, & Thompson, 2016). Hereafter, “wakefulness” therefore refers only to the graded subjective experience of tonic alertness.

Here is the second example. We can expect different degrees of absorption and content superimposition. For example, if the “absence of intentional content” demanded by Low Complexity refers only to a complete absence of mind-wandering and active high-level thinking, then episodes of clear, effortless mindfulness while still perceiving the physical body and visually observing the environment with an entirely quiet mind (as for example in Buddhist walking meditation) already count as an MPE experience. A multimodal, egocentric frame of reference would still be preserved and superimposed onto MPE, like an abstract prior that has lost some but not all of its force. “MPE” then simply refers to the region of phenomenal state-space that we get by subtracting from the natural wake state the following attributes:

a. the phenomenology of cognitive agency (as described in Metzinger, 2013, 2017);

b. the different types of spontaneous, task-unrelated thought (Fox & Christoff, 2018; Seli, Kane, Smallwood, et al., 2018);

c. all episodes in which perception is completely uncoupled from attention (mere “mind-blanking”; cf. Kawagoe, Onoda, & Yamaguchi, 2019; Ward & Wegner, 2013).

On the other hand, if “absence of intentional content” also refers to the content of perceptual, sensorimotor, interoceptive, or emotional states, then we make a stronger demand. If there is no sensory or other representational content whatsoever superimposed onto an MPE experience, then “MPE” refers to a much rarer kind of state. A full satisfaction of the Low-Complexity constraint means minimal complexity of reportable representational content on the level of conscious processing. Therefore, the prototypical region in state-space will be high on the Wakefulness dimension, and minimal in terms of content complexity. From now on, I will call this type a “full-absorption episode.”

9Please note how the phenomenological distinction between prototypical, high-Wakefulness/Low-Complexity states versus the variable spectrum of high-Wakefulness/Low-Complexity states does not cleanly map onto Walter Stace’s conceptual distinction between extrovertive mysticism and introvertive mysticism. Extrovertive mysticism exhibits four clusters of phenomenological features – fundamental unity, certainty about objective reference (i.e., William James’ “noetic quality”, cf. note 2), paradoxicality, and bliss/beatitude/joy – the last three of which do not seem to correspond to the ineffable contentlessness of full-absorption episodes (cf. Stace, 1960, p. 67). Introvertive mysticism is much closer to prototypical MPE, and here Stace also includes all of the features we would today call ego-dissolution (cf. 1960, Chapter 1, section 8), atemporality, and lack of spatial self-location Stace (1960, p. 110) when he aptly describes the phenomenology of an “undifferentiated unity” which holds together the stream of
Case study #4. After the content of consciousness had faded away, suddenly the experience of pure consciousness came. Suddenly I felt that consciousness expanded, boundaries in time and space disappeared, and fullness, wholeness appeared, consciousness was completely without any content. I was completely awake, the awareness was there, very clear, but it was empty of any content, just pure consciousness in itself. When the experience was over I once again became aware of the sounds in the room, and the thoughts began wandering, [...].

(Severeide, 1990, p. 1570; quoted from Bachmann, 2014, p. 52)

Case study #5. I had been meditating alone in my room all morning when someone knocked on my door. I heard the knock perfectly clearly, and upon hearing it I knew that, although there was no “waking up” before hearing the knock, for some indeterminate length of time prior to the knocking I had not been aware of anything in particular. I had been awake but with no content for my consciousness. [...] The experience was [...] utterly without content, [...].

(Forman, 1986, p. 55)

Beyond Wakefulness, the content of a full-absorption episode cannot be reported, because the self-referential mechanisms of forming an autobiographical memory are suspended. A full-absorption episode is a radical disruption of self-consciousness, because MPE has not yet contracted into a first-person perspective. Therefore, only the process of entering into it or of emerging out of it can be faithfully represented in the autobiographical self-model; as such the episode itself is not a part of the subject’s inner life narrative. On the other hand, it may be helpful to distinguish between concurrent ineffability and retrospective report, and we do know from amnesia and trauma research that an absence of reportable conscious experience need not imply an absence of memory.

Before moving on to the third constraint, let us observe how the first two items already allow us to differentiate between three relevant classes of phenomenal states. For simplicity, let us also assume that the Wakefulness constraint is fully satisfied in all three cases:


©The author(s). https://philosophymindscience.org ISSN: 2699-0369
• Maximally satisfied **Low Complexity** plus maximal **Wakefulness** would lead to what I have termed “full-absorption episodes”. In such episodes, the hypothetical phenomenal character of MPE would become a stand-alone feature; “pure awareness” would be the only aspect that could later be reported (see Winter et al., 2020, for a recent single-case study). Empirically plausible examples are “witness consciousness”\(^\text{11}\) during lucid dreamless sleep (for an introduction, see section 5 in Metzinger, 2019) and deep meditative states during the practice of calm abiding and highly focused attention, for example in Buddhist *jhāna* practice.

• Highly satisfied **Low Complexity** would lead to the prediction that at least some representational content remains active, while the stable experience of wakeful awareness presents itself as a continuous and explicit phenomenal quality. Typical examples would be clear and relaxed states of restful alertness without any thought activity (e.g., during classical sitting meditation with closed eyes), or any natural form of mindfulness co-occurring with

\(^\text{11}\)The notion of “witness-consciousness” refers to a term in classical Advaita Vedanta philosophy (*sākṣin*; Sanskrit: साक्षी). It is an example of a phenomenological notion with a long tradition in Indian and Tibetan philosophy, quite obviously bearing direct relevance to the present sketch of an interdisciplinary research program. The earliest appearance of the idea seems to be found in the Atharva Veda, in the *Mundaka Upanisad* (Chatterjee, 1982, p. 339), whereas the source of the philosophical concept *sākṣin* is obscure (Fort, 1984, p. 287, n. 2; Gupta, 1998, p. 19). Witness-consciousness is that which makes all knowledge possible, cannot itself become an object of knowledge, and is self-luminous. Bina Gupta gives the following epistemological characteristics:

1. The witness consciousness, although the basis of all knowing, is different from the object known. It is implied in every act of knowing. It is the ultimate subject; it can never become an object of knowledge.
2. It is the pure element of awareness in all knowing. It is one, immutable, indivisible reality.
3. It shines by its own light; it is self-luminous.
4. It is different from the empirical individual (*jīva*; Sanskrit: जीव), who cognizes and enjoys. In other words, it is different from the empirical individual who is caught up in the triple states of waking, dreaming, and dreamless sleep. (Gupta, 1998, p. 18)

On a purely phenomenological reading that abstracts away from all metaphysical and epistemological aspects of the term “witness-consciousness”, we find something that strongly resembles MPE. It is a self-luminous form of bare awareness which can be superimposed by intentional objects during single “acts of knowing”, a non-egoic type of phenomenal character which has nothing to do with any phenomenal self-representation as a cognitive agent or an entity possessing subjective preferences; and, perhaps most importantly, it is something that could in principle be present in waking states, during dreaming, and in dreamless sleep, simply because it is different from any phenomenal self-model portraying only the “empirical individual”; cf. Fort (1980).

In short, witness-consciousness is an intrinsically non-reactive, non-egoic, and non-conceptual experience of knowing and wakefulness. One interesting reading is of the *sākṣin* as a “field” of consciousness, within which the distinction of subject, object, and the process of knowing arises as a “context for contents, or ‘space’ in which forms may appear” (Fort, 1984, p. 278). For recent discussions and attempts at a definition, see Albahari (2009, p. 66); Fasching (2011, p. 211).

https://doi.org/10.33735/phimisci.2020.I46
©The author(s). https://philosophymindscience.org ISSN: 2699-0369
perceptual experience (e.g., during shamata-practice or “body scan”), or as a natural form of mindfulness co-occurring with bodily motion (e.g., during walking meditation [kinhin], the practice of tai-chi, or the Feldenkrais method).

- Unsatisfied Low Complexity plus maximal Wakefulness could serve to describe so-called “awakened” states. For example, following a global Gestalt switch in which MPE has turned from an unrecognized background into the dominant and spontaneously present foreground of all experience, an explicit experience of Wakefulness could appear as the medium or the centre- and timeless epistemic space in which rich, variable, and complex phenomenal contents unfold over time.

Let us now add another dimension. In many classical texts, the phenomenology of Wakefulness in meditation is described as functionally autonomous, for example as “self-generating ever-fresh awareness” (Mañjuśrimitra et al., 1987, xi) or as originary, naturally present, and non-transient “primordial knowing” (ye shes in Tibetan; cf. Higgins, 2012, p. 447, 2012, p. 99). This phenomenological feature is shared by Self-Luminosity, my third example of a phenomenological constraint that is frequently mentioned in a specific subset of the relevant literature (e.g., in Tibetan Buddhism).

- Self-Luminosity (PC3): a phenomenal property instantiated during some MPE episodes, typically described as “radiance”, “brilliance”, or the “clear light” of primordial awareness.

Phenomenological descriptions of Self-Luminosity are a good example of cultural variance. As opposed to Wakefulness and Low Complexity, Self-Luminosity is not frequently found in Western phenomenological reports. It may therefore not be a necessary element of the common phenomenological denominator we are looking for and may not figure in the final theoretical concept of minimal phenomenal experience. In particular, it could be a metaphorical way of referring to Wakefulness:

Some people compare this cognizance to a radiant “thing” that shines with a light like a “clear light”. It means a sense of being wakeful, a quality of being vividly wide awake, which is empty of any identity and naturally alert. (Rinpoche, 2013b, p. 163)

Let us look at a classical example taken from the Bardo-Thödol, the “Tibetan Book of the Dead” (8th century CE):

**Case study #6.** This brilliant emptiness is the radiant essence of your own awareness. It is beyond substance, beyond characteristics, beyond colour. [...] The instant of your own presence is empty, yet it is not a nihilistic emptiness, but unimpeded radiance, brilliant and
vibrant. [...] Your own awareness, a vast luminous expanse, clarity inseparable from emptiness, is also the Buddha of unchanging light, beyond birth and death. Just to perceive this is enough. If you recognize this brilliant essence of your own awareness as Buddha Nature, then gazing into it is to abide in the state of enlightenment. (Lingpa, 2014, pp. 14–15)

Functional autonomy and a self-revealing “clarity inseparable from awareness” are the first two features we can isolate. Interestingly, there seems to be a third aspect of the relevant phenomenal character that can be described as a concrete, quasi-sensory phenomenology of self-luminosity (sometimes translated as Eigenstrahlkraft in German). How can we make sense of it? I think there are three interrelated ways to do so: as the raw intensity of an internal stimulus source, which, via contemplative practice, has been made introspectively available; as functional autonomy; and as intrinsic epistemicity.

First, I would propose that the relevant phenomenology of quasi-perceptual concreteness consists in the fact that the subjective experience of tonic alertness varies along a dimension of intensity, although it is clearly mode-neutral and therefore not associated with any interoceptive or exteroceptive sensory modality. Something always radiates, but to different degrees. If this first interpretation is valid, then it follows that it should in principle be possible to introspectively trace back the radiance to its origin or hidden cause (Buswell, 1991).

Second, not only does the subjective character of self-luminosity come in variable strengths and degrees, it is functionally autonomous and permeates all of phenomenal space. Therefore, “self-luminosity” may actually be a visual analogy for the experienced functional autonomy of tonic alertness. It may be an attempt to point not so much to a distinct form of phenomenal character or representational content, but to a functional dimension structuring the space of conscious experience, which in turn allows for a relative ordering of global states (Bayne et al., 2016, p. 410). It is also conceivable that an organism may possess an inner repre-

If one takes an intellectually honest, analytical perspective, it is not easy to make sense of passages like these. Prima facie, terms like “brilliant” and “radiant” refer to the sensory phenomenology of brightness, which is a necessary element of all visual experience. “Your own” seems to refer to a phenomenology of ownership, that is, at least a minimal sense of selfhood. “Vibrant” either refers to tactile phenomenology, or – perhaps like the terms “brilliant” and “radiant” – it is used in a purely metaphorical fashion. “Perceiving” is standardly defined as a sensory object representation unfolding under a consciously experienced first-person perspective. Similarly, if “gazing into” is interpreted as a goal-directed process, it must be an active form of visual perception involving an epistemic agent model (an EAM, as introduced in footnote above). The phenomenological configuration described is a highly complex one; it is certainly not “pure” in terms of an absence of intentional content or subject/object structure. Rather, it seems to involve sensory qualities, an active phenomenal self, and a first-person perspective. It would be easy to present dozens of very similar descriptions referring to the phenomenology of self-luminous and “pure” awareness, mostly from the Advaita and Tibetan Buddhist traditions, see Kunsang, Schmidt, & Tweed (2012), Rgyal-ba-g’yung-drung (2017), Namgyal (2006), Namgyal (2019), Rinpoche & Namgyal (2011). Are such descriptions intelligible?


©The author(s). https://philosophymindscience.org ISSN: 2699-0369
sentation of this dimension. Self-Luminosity may therefore be a way of talking about Wakefulness: Wakefulness is always there whenever there is conscious experience at all, autonomous and largely independent of the ever-changing kaleidoscope of surface phenomenology, but it comes in different degrees of intensity and somehow it “radiates.”

Third, Self-Luminosity – if read not as referring to some mysterious form of visual content, but as an epistemological metaphor – also bears a direct relation to Epistemicity, arguably the most important constraint to be extracted in the following section. MPE is a necessary condition for conscious knowledge to be possible, and this includes self-knowledge. In the context of a visual metaphor, for something to be “self-luminous” means that it possesses a quality of intrinsic epistemicity, because it autonomously makes itself knowable. Think of a single burning candle in an unobstructed, dark room. It not only illuminates the room, but also endows itself with visibility. It is not something that needs to be discovered or gradually constructed with epistemic effort, in an agentive, top-down manner. If we take the visual metaphor seriously, MPE must be continuously self-revealing (Strawson, 2017) and transcendental (Higgins, 2013, p. 133) at the same time – a self-intimating internal process that first creates the possibility of knowledge.

2.2 Consciousness as such in Western philosophy of mind

Let us now extract three further semantic constraints for a working concept of “pure consciousness”, but this time from Western analytical philosophy. Below are two passages from a now-classical paper by G. E. Moore, titled “The refutation of idealism”. The second one is a canonical phenomenological description of the relevant kind of introspective attention:

13“Self-luminosity”, under this second interpretation I am here offering, may therefore be best treated as a visual analogy for an amodal type of phenomenal character, as an attempt to express what it is like: A consciously experienced visual quality like “redness” cannot be analyzed merely as hue plus saturation since it also always varies along a third subjective dimension of “brightness”. Phenomenally experienced wakefulness might be a like a domain-general form of brightness which comes in varying degrees of intensity. Zero brightness means zero phenomenal experience. The subjectively experienced brightness of a red object represents sheer stimulus strength for the organism, that is, the luminance of a visually given physical object – and perhaps MPE literally turns out to be a mode-neutral form of brightness. For example, the internal “luminance” in question could be a dynamical property of some part of the neural body. Attending to “luminous emptiness” might really be attending to the autonomous activity of this body part, and it might even be the case that, as a causal consequence of additional representational resources, the local body part begins to reflexively represent itself (for an excellent discussion see Bachmann, 2014, p. 139). Therefore, if we want to understand the phenomenal concreteness of mode-neutral and contentless wakefulness, then we have to ask: What, for bare wakefulness, is the non-sensory equivalent of the raw stimulus intensity caused by the physical property of luminance in a given visual target and which, in the domain of visual consciousness, leads to the experiential quality called “brightness”? Is there something like “mode-neutral brightness”, and can it be “made reflexive” by introspective attention? What exactly is the stimulus source, its hidden cause? If we read “self-luminosity” as a visual analogy, then the question becomes: What exactly is “luminant” here, and is it perhaps a part of the neural body itself?
when we refer to introspection and try to discover what the sensation of blue is, it is very easy to suppose that we have before us only a single term. The term “blue” is easy enough to distinguish, but the other element which I have called “consciousness” – that which a sensation of blue has in common with a sensation of green – is extremely difficult to fix. [...] And in general, that which makes the sensation of blue a mental fact seems to escape us; it seems, if I may use a metaphor, to be transparent – we look through it and see nothing but the blue; we may be convinced that there is something, but what it is no philosopher, I think, has yet clearly recognised. (Moore, 1903, p. 446)

Case study #7. [...] the moment we try to fix our attention upon consciousness and to see what, distinctly, it is, it seems to vanish: it seems as if we had before us a mere emptiness. When we try to introspect the sensation of blue, all we can see is the blue: the other element is as if it were diaphanous. Yet it can be distinguished if we look attentively enough, and if we know that there is something to look for. (Moore, 1903, p. 450)

First, it is noteworthy that Moore actually uses the term “emptiness”, which plays a great role in Buddhist philosophy (cf. the preceding quotation). Second, it is noteworthy that he also uses a visual metaphor to describe the mode-neutral phenomenal character in question. His paper is the locus classicus for the concept of “phenomenal transparency”, a property of some conscious states which he called “transparency” or “diaphanousness”. Today, a standard definition of “phenomenal transparency” is that it essentially consists in only the content properties of a conscious mental representation being available for introspection, but not the fact that it also has non-intentional or “vehicle properties”. In other words, it is not experienced as a representation. Introspectively, we can access its content, but not the content-formation process itself. As I have extensively argued elsewhere (Metzinger, 2003a, 2003b, 2014), the empirically most plausible explanation for this fact is that, relative to the given temporal resolution of introspective attention, earlier processing stages are invisible to the system, because all competing interpretations have already been disambiguated. In philosophy, it is often assumed that transparency in the sense of the standard definition above is a property of all phenomenal states. But of course, the standard assumption is incomplete: opaque phenomenal representations also exist (whereas unconscious states are neither transparent nor opaque in this sense). The most notable phenomenological examples of opaque state-classes are consciously experienced thoughts: We experience them as mind-dependent and internally constructed, as mental representations that could be true or false. Apart from that, some types of emotions, pseudo-hallucinations, and lucid dreams are also subjectively experienced as representational processes. They make the possibility of misrepresentation introspectively available, because the associated subjective confidence changes over time. A spectrum therefore exists between phenomenal opacity and phenomenal transparency, and any given
content can vary along this spectrum. Phenomenally opaque processes sometimes appear to us as deliberately initiated cognitive or representational processes. However, they can also appear as automatic or spontaneously occurring, as limited or even global phenomenal simulations; they can draw us into a manifest daydream; and they frequently seem not to be under the experiential subject’s control (for more, see Metzinger, 2003a, 2003b, 2014; Pagnoni, 2019, p. 308).

Moore presents a phenomenological argument against the transparency of consciousness as such. Many authors have read Moore as if he were talking about the “transparency of qualia.” However, the relevant point concerns not the transparency of awareness or qualia, but rather the fact that consciousness itself can be made phenomenally opaque. Moore’s much deeper point is that conscious experience as such has a unique phenomenal character sui generis (namely, a second-order “awareness-of”, a relational “signature of knowing”), and that this character can sometimes be detected by introspective attention. In fact, Moore’s self-stated goal in introducing “transparency” is actually “to try to make the reader see it” (Moore, 1903, p. 450). His important phenomenological claims are the following:

- **(M1)** There is one most general phenomenal property, which is shared by all sensory qualities:
  → “Consciousness” as such, pure awareness, phenomenality per se.

- **(M2)** Under standard conditions, this global property is “transparent” (Moore, 1903, p. 446), a “mere emptiness” or “diaphanous”:
  → It is usually not explicitly experienced, but it can be.

- **(M3)** This property is evasive:
  → It “seems to vanish” under attentional agency; i.e., if we actively try to “fix our attention” on it (Moore, 1903, p. 450), thereby creating an epistemic agent model (see footnote 7).

- **(M4)** This property can become phenomenally opaque, under two conditions:
  → we look “attentively enough”,
  → and we “know” that there is a possible object for introspective attention (Moore, 1903, p. 450).

- **(M5)** It is difficult to simultaneously direct and sustain introspective attention to the global property in question and to concrete perceptual qualities:
  → Moore found no philosopher in the literature who was “able to hold it and blue before their minds and to compare them, in the same way in which they can compare blue and green” (Moore, 1903, p. 450).

---

14 For an excellent, well-researched, and substantial discussion containing further references, see Hellie (2007).
• **(M6)** Awareness is relational and consciousness is a second-order epistemic process:

→ Having a sensation is an awareness of something, and consciousness is the knowledge that this awareness currently exists. In his own words:

The true analysis of a sensation or idea is as follows. The element that is common to them all, and which I have called “consciousness”, really is consciousness. A sensation is, in reality, a case of “knowing” or “being aware of” or “experiencing” something. When we know that the sensation of blue exists, the fact we know is that there exists an awareness of blue. (Moore, 1903, p. 449)

• **(M7)** Consciousness is meta-awareness:

→ The type of epistemic relation is identical: “awareness-of” is the same relation in sensation as it is in becoming aware of this awareness. Once again, in Moore’s own words:

This relation is just that which we mean in every case by “knowing”. To have in your mind “knowledge” of blue, is not to have in your mind a “thing” or “image” of which blue is the content. To be aware of the sensation of blue is not to be aware of a mental image of a “thing,” of which “blue” and some other element are constituent parts in the same sense in which blue and glass are constituents of a blue bead. It is to be aware of an awareness of blue; awareness being used, in both cases, in exactly the same sense. (Moore, 1903, p. 449)

Given this analysis, one can now interpret Moore as saying that a) an aperspectival and non-agentive form of meta-awareness is co-instantiated with all forms of conscious perceptual knowledge; that b) the second-order epistemic relation of an awareness-of the current existence of an awareness-of some specific perceptual phenomenal character is mostly unnoticed (i.e., like a mind-independent and transparent medium in which this character simply “appears”); but that c) we can make it phenomenally opaque by attending to it in the right way. From this interpretation we can extract three further phenomenological constraints to semantically enrich our six-dimensional working concept of MPE:

• **Introspective availability (PC4):** We can sometimes actively direct introspective attention to consciousness as such and we can distinguish possible states by the degree of actually ongoing access.

• **Epistemicity (PC5):** The phenomenal experience of knowing, which comes in degrees and can also be described as the subjective quality of confidence.15

15To avoid any misunderstandings at this point: The phenomenology of knowing can exist without
• **TRANSPARENCY/OPACITY (PC6):** Like all other phenomenal representations, MPE can vary along a spectrum of opacity and transparency.

Constraints 4 – 6 can now be interestingly discussed in the context of the first three constraints presented in the previous section. Here are some examples.

### 2.2.1 Introspective availability

**INTROSPECTIVE AVAILABILITY** has been endorsed not only by G. E. Moore, but by many Eastern contemplative traditions. However, it is important to always clearly distinguish between the dispositional property of *availability* for introspective attention and actually ongoing introspective access to the relevant phenomenal character. “Availability for introspective attention” is a functional property a given internal state may or may not have, whereas ongoing introspective access presupposes the capacity for active attentional control, at least at the beginning of an episode. This actually ongoing access may then change the overall state. At the same time, it is also clear that the phenomenology of ongoing introspective access can only be a transitory phase, and that it is not itself an element of the relevant conscious state, which is devoid of subject/object structure, mental agency, and any form of conceptual or propositional mental content. Here is how the relevant transition is described from a Tibetan Buddhist perspective:

**Case study #8.** The moment you recognize, it is already seen. There is nothing extra remaining that you missed. This is not like space looking at itself, because space does not see anything. When your mind, which is cognizant, recognizes itself, you immediately see that there is no “thing” to see. It is already seen in the same moment. At that very moment there is no thought, because the present thought has naturally vanished. (Rinpoche, 2013a, p. 79)

In the following description of contemplative phenomenology, **Wakefulness** and **Epistemicity** are interpreted as aspects of a self-sustaining, functionally autonomous process that precedes introspective access:

there being knowledge in the epistemological sense. PC5 is an exclusively phenomenological constraint, and it can be satisfied by “noetic” experiences (cf. footnote 2) as well as intuitions, hallucinations, illusions of control, complex delusional states, and all other forms of consciously experienced misrepresentation that are not experienced as misrepresentations.

The concept of “introspective attention” is only indirectly related to classical theories about “metacognition” (Flavell, 1979; Metcalfe & Shimamura, 1996). Introspective attention is non-conceptual, a subsymbolic process that has nothing to do with “beliefs”, “thinking about thinking”, or declarative knowledge about one’s own states. Rather, it is related to the optimization of precision expectations relative to internal stimuli, cf. Wiese and Metzinger (2017), Pagnoni (2019). In terms of contemplative practice, the goal will be an effortlessly sustained, entirely non-conceptual and non-propositional state of meta-awareness. For a useful conceptual clarification, see Dunne, Thompson, and Schooler (2019).
Case study #9. Empty cognizance is an original state that is never lost. It merely needs to be known. We need the knowledge of self-existing empty cognizance – the knowing itself to be empty and cognizant. (Rinpoche, 2013b, p. 114)

There seem to be three elements here: first, a functionally autonomous state of bare subjective confidence, of “knowing without an object”; second, a transient episode of introspective attention directed at this state; and third, the emergence of a new phenomenal state which dynamically integrates both elements (see also Rinpoche, 1995, pp. 106–129). If MPE were to be the first element – in terms of minimally complex, full-absorption states characterized by a full satisfaction of the Low-Complexity constraint (as discussed above) – then it would be a functionally autonomous kind of internal process that could in principle be actively introspected, but mostly is not.

Let us now assume the process in question would be the process by which the central nervous system activates itself, creating the property of tonic alertness (Metzinger, 2019, sec. 6). If attended to, the original process could be “made reflexive” or “ignited” into a recurrent loop, thereby causing a new dynamical state to appear – plausibly yielding a distinct, new type of phenomenal character that we might provisionally call “self-aware wakefulness”. Compare this classical phenomenological text from the 16th century:

Case study #10. […] it is an aware emptiness that defies any description of being such-and-such – it is inexpressible and yet it can be experienced. […]
It is a wakefulness for which no words suffice. It is not a definable entity, but at the same time, it is a self-knowing aware emptiness that is clear, lucid, and awake. (Namgyal, 2001, p. 29)

This demonstrates an important difference. For Moore, concrete introspective access is mere meta-representation that creates phenomenal opacity by making a global and previously transparent aspect introspectively available for a robust conscious self. However, contemplative traditions clearly favour not higher-order representation, but a first-order, reflexive theory of meta-awareness (Ramm, 2019, p. 7). Here, we find the notion of non-egoic reflexivity: a minimally complex, func-

---

17For this reason, many contemplative traditions (e.g. rDzogs chen) stress that the relevant state cannot be directly introspected because it is a self-constituting process which is functionally independent of all possible mental efforts a meditator could ever make, and because any such effort would a) contract the overall experience into an introspective first-person perspective (the “meditative self”) and b) immediately impose a subject/object prior onto it. In Western analytical philosophy of mind, see for example Michelle Montague for the point that awareness of awareness cannot be transformed into direct introspective awareness and that it cannot be directly introspected; see Montague (2017, p. 22), and for the notion of a “self-constituting representational relation”, see Montague (2016, p. 59). In 2009 Uriah Kriegel first developed this idea into the “attention-shift model of introspection”, which points out that some phenomenological elements like “peripheral awareness” may well be manifest without being introspectable. Part of the
tionally autonomous baseline state that dynamically represents itself on a subsymbolic level (Josipovic, 2019, sec. 3). Perhaps we could metaphorically describe the phenomenological correlate as “an emptiness that has awoken to itself”.

On the other hand, the dispositional property of Introspective Availability itself can be interpreted as a second functional dimension (as opposed to a content-related one). Nevertheless, it has to do with what Bayne, Hohwy, and Owen (2016, p. 409) call a “gating” of the contents of consciousness. MPE itself may be a very simple, highly abstract, and unstructured type of phenomenal character, but there is variability along a dimension of actually ongoing introspective access. It might open the organism to other kinds of conscious contents, for example by making it sensitive to incoming perceptual stimuli or enabling orientation in time and space. Adding introspective attention would transiently create a sense of effort, a phenomenal self, and an internal, self-directed first-person perspective. This could then be followed by brief periods of effortless meta-awareness (i.e., the non-egoic but reflexive form of “empty cognizance” described above) in which the first-person perspective briefly dissolves until the next train of apparently spontaneous, task-unrelated thought ensues. Interestingly, this emerging picture is in full accordance with recent models of what actually happens during the process of mindfulness meditation (Hasenkamp, 2018; Hasenkamp, Wilson-Mendenhall, Duncan, & Barsalou, 2012).

### 2.2.2 Epistemicity

Epistemicity can now be read as a prediction about the phenomenal character of more complex episodes in which introspective attention has been transiently added to a simple state of MPE. Attentional agency creates the phenomenology of a knowing self, an epistemic agent model (Metzinger, 2017; see also Limanowski & Friston, 2020, this issue), and such states would then still be characterized by an explicit subject/object structure. Here, however, the additional idea is that there may be a non-agentive variety of “knowingness” in low-complexity MPE states lacking this structural feature but characterized by a more globalized and effortless form of meta-awareness. Contemplative practice ultimately dissolves itself and reveals the phenomenology of subject/object structure to be merely “parasitic” (Higgins, 2012).
on an entirely non-dual and more fundamental level of experience. After attentional agency ends, a distributed continuum of subjective confidence remains. The prediction would be that this should generate the phenomenal character of an all-pervading field of “distributed confidence” or “primordial knowing”. Recall that Moore introduced “awareness” as an epistemic process (M6, M7). In contemplative phenomenology, the specific phenomenal character of Epistemic-ity I am here trying to isolate has sometimes been termed “the flash of knowing that gives awareness its quality” (Mañjuśrimitra et al., 1987, xxiii) and it is often described as a content-independent “cognizant quality”:

**Case study #11.** In that moment, we don’t see any concrete thing whatsoever. There is an immediate knowing that the essence is empty. There is something that cognizes that the mind is empty, and this cognizant quality is indivisible from the emptiness itself. At the moment that this is an actuality, you don’t need to do anything more. (Rinpoche, 2013a, p. 89)

MPE is not a mystical experience, but in research on mystical experience, this phenomenological feature has often been described as a unitary state exhibiting a “noetic” phenomenal quality, while still lacking any specific form of content (Wainwright, 1981). In such states one would predict a phenomenal signature of knowing which a) has no object (i.e., a form of knowing or subjective confidence without content), and which b) has been globalized (i.e., which is not object-oriented or tied to a localized model of a knowing self any more). Here is one phenomenological description from an advanced meditator:

**Case study #12.** “[…] there’s no personal point of view, it’s the world point of view, it’s like the world looking, not ME looking, the world is looking.” (quoted from Ataria, Dor-Ziderman, & Berkovich-Ohana, 2015, p. 141)

Compare this with the following, more complex phenomenological analysis:

**Case study #13.** The first characteristic is that it has no boundaries, no fence round it, no edges; it’s absolutely unlimited in all directions. The second characteristic is that it is absolutely clear, clean, empty of contamination. It is utterly simple, totally transparent, empty of everything but itself, empty even of itself, clearer than glass, cloudless, an infinite sky. The third characteristic is that it is also full of the world. Because it’s empty, it’s full – full of the scene, whatever the scene is, absolutely united with it. The fourth characteristic is that it is awake, it’s aware, it’s conscious. And the fifth characteristic is that it is right where you are. (Harding, 2000, p. 65)

Clearly, MPE is neither an internal model of a self-as-subject, nor a model of an object. Philosophically, this creates a conceptual tension: If we want to follow

---

For more on the relevance of Douglas Harding, see Ramm (2019).
Moore in interpreting the concept of “awareness” as an epistemic relation, then it becomes difficult to specify the subject-argument place. What exactly are the relata of this special relation? What is the known, and what is the knower? What is the “something that cognizes” in case study #11, and what do we make of the fifth characteristic in case study #13? It may turn out that the experience itself isn’t really contentless, but only described as such, and it may still be non-egoic in the sense that no knowing self (i.e., no epistemic agent model) is part of its content. But, under an epistemological analysis, who or what is the epistemic subject?

I want to submit that the subject simply is the organism or conscious system as a whole, and that MPE could be an internal model of an epistemic space currently opened by it – and that the as-yet-unpartitioned nature of this space is what explains the abstract phenomenal character of epistemicity. Call this the “ESM”-thesis (ESM for “epistemic space model”): MPE is an integrated inner representation of a space of possible knowledge states or epistemic policies, defined by the organism’s set of epistemic capacities. Consciousness would then appear whenever an information-processing system has a) created an internal epistemic space, and b) activated a model of this very epistemic space within it (Metzinger, 2019, sec. 8.4.1). Using Moore’s terminology, we could say that typically this model is “diaphanous” or “transparent.” Just like physical space, epistemic space creates an allocentric frame of reference, it is unbounded and non-centred, while of course its variable content is perspectival: being a simulation of an embodied behavioural space, it is egocentric, there are perceptual horizons, and so on. However, just as physical space cannot be touched, felt, or grasped, MPE as such cannot be “touched by attention” or “grasped by thought”, because this would only introduce subject/object structure into a phenomenal baseline state, in which we have already been immersed all along. In my view, the best phenomenological metaphor for “pure consciousness” or MPE, to be found in many places in the contemplative literature, is the epistemic-space metaphor:

Space is the example for mind essence, because space is unmade. But mind essence is not totally like space, in that space cannot think. Space has no knowing. Our mind is cognizant emptiness – empty like space, but with a natural knowing. That union of cognition and emptiness is seen when recognizing. It is immediate, like the example I mentioned of pointing into mid-air. You do not have to wait to raise your arm for your finger to touch space – you are already touching space, all the time. You do not have to move your hand forward; the contact

---

By definition, an epistemic space is a space of possibilities, it contains every possible epistemic scenario and every dynamic partitioning of itself that could ever take place, everything that could potentially be known and experienced by a given system. Therefore, the internal model of such a space, of the ESM as such, would be characterized by a very high degree of counterfactual invariance: Whatever the system could know or experience would necessarily take place within it. Here, “maximal counterfactual invariance” (for the ESM itself) and “maximal counterfactual richness” (for its content) could be an excellent way to analyze the Moorean concept of “transparency”.


©The author(s). https://philosophymindscience.org ISSN: 2699-0369
is already occurring and has been your entire life. All you have to do is recognize that it is taking place. It’s the same with mind essence. (Rinpoche, 2013b, p. 112)

If Moore is right, then MPE actually is an epistemic relation (“awareness-of”). If MPE is successfully introspected (i.e., without – as Moore and the Tibetan rdzogs chen-scholars of the 8th to 14th centuries tell us – trying to “fix” attention on it in an effortful, agentive way), then it is natural to expect a distinct phenomenal character of knowing. However, in this very special case there would be no correlated epistemic object and all the other negative constraints listed in section 2 would be simultaneously satisfied as well. “Epistemicity” therefore refers to the mere phenomenal signature of knowing, i.e., a functionally autonomous experience of subjective confidence or certitude, a representation of model evidence as such. MPE could then be the special phenomenal character of non-directed, non-agentive and aperspectival “knowingness” – but with or without a perceptual object, depending on the degree of absorption discussed above.

An additional phenomenological aspect of PC5, the epistemicity constraint, is that MPE states are sometimes described as reflexive, and sometimes not. That is, while awareness as such has no external intentional object it is sometimes described as having itself as an object, as being self-reflexive in a way that is non-conceptual and which does not necessarily involve the experience of a substantial, separate self (MacKenzie, 2008).

If there were such an intrinsic and fundamental form of reflexivity, then one could imagine it either as co-existing with specific content or as a pure stand-alone phenomenon. For example, it would be conceivable that it co-exists with complex content during the daytime while normally remaining as entirely unrecognized by the subject of experience. At the same time one can now make sense of the idea of how it might occur in dreamless deep sleep, namely, as an “island of awareness” the content of which “is not shaped by sensory input from either the external world or the body and which cannot be expressed via motor output” (Bayne, Seth, & Massimini, 2013).

This is interesting, because there also seems to be a graded but distinct phenomenal character of “certitude” or subjectively experienced confidence, which we can now analyse as a non-conceptual representation of the “degree of belief”, “plausibility”, “trustworthiness”, or “reliability” assigned to an inner or outer perception, or any other consciously experienced epistemic state. We could also term it the “phenomenology of validity” (or the “noetic quality” mentioned in footnote 2), and it is clearly an essential aspect of MPE. It can be found in complex hallucinations, various psychiatric syndromes, or during epileptic seizures as a result of cortical dysfunction; cf. Picard (2013). What makes the phenomenal character of subjective confidence especially interesting from a philosophical point of view is a) that it relates directly to centuries of theorizing about consciousness as “knowing that one knows”, and b) that there are now unifying conceptual frameworks describing confidence as Bayesian probability, bridging different fields of psychology and neuroscience and unifying disparate research perspectives. I will come back to this point in later publications, see also Meyniel, Sigman, & Mainen (2015). Epistemicity could be the way Bayesian probabilities and model evidence are represented internally, by the system itself and on the level of conscious processing. MPE is a highly relevant target for research, because a better understanding of Epistemicity may help us to study this aspect in isolation.
The possibility of a fundamental, non-egoic form of reflexivity built into consciousness itself has been discussed for centuries by Buddhist philosophers (Finnigan, 2018; Williams, 1998). Following Brentano’s classical idea of a single state simultaneously directed at an intentional object and itself (Brentano, 1874/1973, p. 153, 2012, p. 25), many fine-grained conceptual models have been developed in more recent analytical philosophy of mind (Montague, 2016, Chapter 3, 2017; Peters, 2013), and the issue is currently attracting a lot of attention (Ganeri, 2017; Josipovic, 2019; Kriegel, 2019; Strawson, 2017; Thompson, 2011). Please note how I am here only interested in the phenomenal character of intrinsic reflexivity itself, in the reported experience of an explicit, introspectively available, non-egoic, non-conceptual, and exclusively self-directed awareness. This character is described in the following case study:

**Case study #14.** It’s absolutely alive to itself, fully conscious, and it’s conscious of itself as No-thing, colorless, featureless, totally empty and totally full. The Space is conscious of itself at the near end, but it has no personal characteristics. (Harding, 2000, p. 20)

What makes epistemicity even more interesting – and potentially relates it to transparency/opacity – is the additional phenomenological fact that, while low-complexity MPE states are often described as states of “pure knowing”, apparently very similar configurations are sometimes also later reported as conscious episodes of “pure being”. What is the relationship between “pure knowing” and “pure being”, if once again we strip away all epistemological and metaphysical interpretations of the relevant reports, staying with just the phenomenal character itself? There obviously exists a globalized, non-agentive phenomenology of epistemicity as well as a phenomenological counterpart of pure onticity. The first type of phenomenal character (i.e., globalized epistemicity without an object, often described as “empty cognizance”) can already be found in the first quotation presented above, case study #11. Let us look at two examples reporting the experience of a globalized and non-dual form of “empty onticity”, of mere existence as such, the first one coming from William James, the “father” of American psychology himself:

**Case study #15.** During the syncope, there is absolute psychic annihilation, the absence of all consciousness; then at the beginning of coming to, one has at a certain moment a vague, limitless, infinite feeling – a sense of existence in general without the least trace of a distinction between the me and the not-me. (James, 1890/1950, p. 273)

Here is another report of the process of regaining consciousness following a severe accident, as an illustration of what I mean by the “phenomenology of onticity”, again without any form of subject/object dichotomy:

**Case study #16.** There was something, and the *something* was not the nothing. The nearest label for the *something* might possibly be “awareness,” but that could be misleading, since any awareness I’d ever had before the accident was my awareness, my awareness of one thing or another. […] In contrast, this *something*, if it be called awareness, had no I as its subject and no content as its object. It just was. (Sullivan, 1995, p. 53)

This is puzzling and interesting at the same time. There seem to be borderline phenomenologies connecting the region of “pure awareness” with that of “pure being” in phenomenal state-space, linking the experience of knowing as such with the experience of being as such. One logical possi-
2.2.3 Transparency/Opacity

Finally, Transparency/Opacity shows why MPE is such an interesting limit case for theories of consciousness. Please recall that a full satisfaction of Low Complexity demanded a complete absence of intentional content for MPE, and that “transparency” was defined on the representational level of description (Metzinger, 2003b). On this level of analysis, transparency is a property of phenomenal representations, and is brought about by the fact that only their intentional or content properties are available for introspection, not the properties instantiated by the dynamics of earlier processing stages (i.e., of what often is called the “vehicle” in traditional philosophical terminology). From this, it follows that for the special case of a putatively contentless state like MPE, if it really were transparent, there would be nothing introspective attention could ever discover – there would simply be no intentional properties that could ever be available for introspection.

If, however, MPE were phenomenally opaque (if it were sometimes experienced as a representation), then we would be forced to conclude that PC1 (Wakefulness) and PC3 (Self-Luminosity) actually refer to an introspective experience of certain properties of the physical carrier itself. For example, such non-intentional “vehicle properties” could simply be earlier processing stages in the brain, certain dynamic aspects of the very construction process that creates the level of phenomenal representation in the first place. As we actually have numerous reports of MPE-like states across the centuries and across cultures, it is now plausible to draw a second conclusion: At least some of these reports may actually be reports about an entirely non-intentional process, not merely a mode-neutral one (Josipovic, 2019).

Now imagine that you would make a previously transparent episode of MPE phenomenally opaque by just the right kind of “looking attentively enough” while at the same time not trying to agentively “fix” attention on it, exactly in the way Moore described. A third conclusion is that if the fundamentality constraint is satisfied and MPE really underlies, envelopes, or permeates all other forms of conscious content, then – if it were made opaque – we would expect a global phenomenal effect. In richly superimposed states involving perception of body and environment everything should then begin to be experienced as made out of pure consciousness, as merely representational, as a virtual model lacking an inherent, mind-independent existence of its own – not as a world, but as the content of an epistemic space which is now explicitly co-represented (cf. case study #12).

bility is that “epistemicity” as well as “onticity” could refer to almost the same kind of global state. This would be again be characterized by the absence of discursive thought, perceptual content, spatiotemporal self-location, agency, etc. The only difference would be that in the first category this global state is still experienced as representational (i.e., as a contentless form of awareness-of) while in the second case it is not. Perhaps the first type of state is reflexive (and therefore experienced as self-representational), while the second one is not. In other words, the relevant difference would then be the difference between a phenomenally opaque and a phenomenally transparent representation of the inner process creating tonic alertness.
In Low-Complexity situations this would lead to the phenomenology of “empty cognizance” discussed above. If, however, there were rich superimposed content while the internal model of our epistemic space were transparent, then this content would simply appear as “real” and directly given, as the world we live in. Looking at MPE in this way begins to yield a theory of consciousness: “Appearance” results from a transparent representation of meta-confidence. Phenomenally experienced “realness” is a transparent representation of certainty or high Bayesian probability – the globalized but currently invisible knowledge that we know.

In Low-Complexity situations the prediction is an experience of “pure being” (see footnote 21). In richly superimposed states, a world would appear. Wakefulness is the phenomenally opaque counterpart to the transparent, conscious experience of “realness”. In full-absorption episodes there would be neither introspection nor reportability, just a first-order model of the non-intentional, physical carrier as such. In absorbed but weakly superimposed states already permitting introspective access the phenomenological prediction is that the phenomenal quality of Epistemicity (PC5) would be co-instantiated with Wakefulness (PC1) and Self-Luminosity (PC3) – an experience of subjective confidence related to bare wakefulness and the functionally autonomous “mode-neutral brightness” discussed above.

3 The difference between phasic and tonic alertness

3.1 Terminology first: Alertness vs. wakefulness

The empirical literature is often unclear and ambiguous regarding the use of notions like “arousal”, “vigilance”, “sustained attention”, “wakefulness”, and “alertness”. However, there are two semantic elements present in a large majority of scientific treatments: first, epistemic capacity, and second, absence or nonspecificity of representational content. For example, Oken, Salinsky, and Elas speak of “activation states of cerebral cortex that impact the ability to process information where the activation itself contains no specific information” (Oken, Salinsky, & Elsas, 2006, p. 1885). It is exactly the representation of such states which is relevant for MPE, but viewed from a philosophical perspective, the existing terminology to describe such states is highly inconsistent. As Oken and colleagues write:

Unfortunately, no terms are ideal to describe these states of cortical activation since most terms are in broad use with varied associations and there are not [sic] perfect physiological markers. The term vigilance, in particular, has unfortunately been used in different ways by different groups of scientists. [...] The field has been hindered by inconsistent or poorly defined terminology. Researchers should be particularly careful about the usage.
of the term vigilance, understanding there are factors contributing to vigilance other than just arousal on the sleep–wake dimension. Avoidance of the term vigilance because of its varied definitions would be most helpful. The use of the term tonic alertness may be preferable since it does not already have a confusing history. (Oken et al., 2006, pp. 1885, 1991)

For the purposes of this paper, let us clearly distinguish between the physical, the functional, and the phenomenological levels of description. I will use the following definitions:

“Arousal”:

- A graded **physical property** of the human brain.
  → A purely physical, but variable boundary condition determining the depth of cortical information processing available to the organism as a whole.
  → A vital parameter that must be predictively controlled (e.g., to generate the sleep–wake cycle).

“Tonic alertness”:

- A graded **functional property** determining the capacity for sustained attention.
  → A causal function resulting from the successful control of arousal over longer periods of time, in the absence of an external cue.
  → A functional property causally enabling cognitive capacities like orientation, executive control, attention, and epistemic agency on the mental level.

“Wakefulness”:

- A graded **phenomenal property** which is sometimes introspectively accessible.
  → A representation of tonic alertness.
  → The major component of the phenomenal character of MPE (PC1).
  → The primary dimension of phenomenal state space.

What is the relationship between these three properties? In humans, the functional property of tonic alertness is realized by certain physical properties of the brain, and we do not yet know what exactly these properties are. Conceivably, it could be realized by very different physical properties in machines that create an integrated internal model of their own epistemic space as such. For neurotypical humans, it is highly plausible to assume that a critical level of cortical arousal is the most central...
necessary condition. Whether the right level of arousal as such is sufficient for tonic alertness is currently unknown, but unlikely.

The phenomenal experience of wakefulness is a representation or, plausibly, a predictive Bayesian model of the functional property of tonic alertness. Both the phenomenal property of wakefulness and the functional property of tonic alertness cannot exist without the brain. But tonic alertness, if it is not represented or internally modelled, can exist without the conscious experience of wakefulness (as in some cases of unresponsive wakefulness syndrome; cf. Laureys et al., 2010). Put differently, an organism can be tonically alert without knowing that it is alert: Consciousness is knowing that one is alert. An organism can embody a rich space of epistemic capacities without having an internal model of this fact.

What do we know about the functional property connecting the physical and the phenomenological levels? Tonic alertness is a global functional property, in part physically realized by the level of cortical arousal, and it is hypothesized to be one of the core functions of the cingulo-opercular (CO) network (Sadaghiani & D’Esposito, 2015). It is one component of what we call the capacity for “attention” in folk-psychological contexts, and its function for the organism consists in “achieving and maintaining a state of high sensitivity to incoming stimuli” (Posner, 2008, p. 193). Our neuroscientific understanding of the CO network is gradually improving, and it already provides us with important pointers that are consistent with our phenomenological analysis: Tonic alertness is an a) sustained and b) internally initiated “preparedness to process and respond” which c) implies a capacity for attentional agency, of being able to “co-engage” phasic alertness and selective attention directed at a specific task (Sadaghiani & D’Esposito, 2015, p. 2763). Just like the quality of mindfulness cultivated by contemplative traditions, it can be described as a global and general mechanism “of keeping cognitive faculties available for current processing demands and holding unwanted activity at bay” (Sadaghiani & D’Esposito, 2015, p. 2764).

It is also empirically plausible to assume that tonic alertness is a central causally enabling factor in the realization of two other functional properties which are distinct but closely related: orientation and executive control. Here, it is important to note that all three functions – alertness as such, orientation, and executive control in terms of attentional agency – are epistemic capacities which, at a given point in time, a biological organism may have or not have. They are the foundation for what Blanke and Metzinger called a “strong first-person perspective” (Blanke & Metzinger, 2009, p. 7).

Please also note how the paradigmatic cases of “pure” consciousness discussed above are precisely characterized by an absence of spatiotemporal self-location

---


©The author(s). https://philosophymindscience.org ISSN: 2699-0369
(i.e., during full-absorption episodes, the subject is not oriented towards time or space) and absence of attentional or cognitive agency (i.e., mental actions like controlling the focus of attention or one’s thought processes). The current proposal says that phenomenality (i.e., a minimal form of conscious experience) emerges precisely whenever the first functional property – tonic alertness, the causal enabler of orientation and executive control – is probabilistically represented. In other words: Minimal phenomenal experience is a predictive model of a specific set of epistemic capacities and the representational space opened up by them, for example, “high stimulus-readiness”, “orientation readiness”, plus the mere “capacity for mental self-control” on the level of attention and cognition. However, it is a temporally shallow model as it accumulates evidence only on a very short timescale and does not entertain complex state transitions or sequences over time (Friston, 2018; Friston, Rosch, Parr, Price, & Bowman, 2017, p. 486). This point is in full keeping with the Low-Complexity constraint and the well-known fact that meditators try precisely to end any sort of inner narrative, by continuously bringing their attention back to the fine details of the present moment, thereby collapsing the deep temporal structure of their model of reality. At the same time it casts doubt on attempts to frame consciousness itself (instead of self-consciousness) as directly related to temporal thickness (Friston, 2018; please also see the contribution by George Deane, 2020, in this special issue). Interestingly, if viewed as a running Bayesian model, MPE has a certain degree of model evidence. If this objective property is introspectively available, then this will generate Epistemicity, the subjectively experienced confidence that the capacities mentioned above will actually lead to the creation of internal states carrying epistemic value in the future. More simply, we could say that MPE is an expectation of knowledge. If this expectation is transparent, it necessarily creates appearance: the certainty that something knowable exists.

3.2 The phenomenology of phasic vs. tonic alertness

We can now formulate the empirical hypothesis that what satisfies Wakefulness and Epistemicity on the phenomenological level is the existence of a high-probability Bayesian representation of tonic alertness. However, on the functional level of description, it is important to first distinguish tonic alertness itself, which is long-lasting and independent of external cues, from short episodes of heightened but externally triggered arousal. The functional difference between these two different kinds of alertness is philosophically relevant, because it generates two fundamentally distinct kinds of phenomenal character.

A paradigmatic example is the rapid change in attention due to a brief event, like an unexpected sound in your environment. The phenomenology of phasic alertness is the experience of being forced to take a new attentional perspective, of re-orientating the “arrow of intentionality” (Metzinger, 2006, 2017). It is a brief episode in which the first-person perspective as consciously experienced is shifted
to a new epistemic target, in a bottom-up, saliency-driven way. Scientifically, phasic alertness is typically studied with infrequent, unpredictable warning signals preceding the presentation of a stimulus, shortening reaction time. Whereas intrinsic, tonic alertness is the basis for operations such as orientating and selective attention, these operations themselves have a distinct phenomenal character of their own: the dynamic process of relocating attentional resources or suddenly shifting to a new target for executive control, typically following an experience of unexpected salience detection. What makes an experience salient depends on many different factors: Perceptual features like contrast, clarity, unexpected size or intensity clearly play a causal role, but the overall emotional, motivational, or cognitive context may be equally important in triggering the phenomenology of phasic alertness. In the words of Weinbach and Henik:

Alertness is commonly divided into two different modes of function: tonic and phasic alertness. Tonic alertness, which is also known as “intrinsic alertness,” designates the internal control of wakefulness or arousal in the absence of an external cue in a top-down manner (Sturm et al., 1999; Sturm & Willmes, 2001). Phasic alertness, on the other hand, is a short-lived effect of achieving high levels of alertness following a salient external event (Posner, 1978, 2008; Sturm & Willmes, 2001). (Weinbach & Henik, 2012, p. 1530)

The phenomenology of tonic alertness is the non-dual conscious experience of intrinsic and sustained wakefulness. It fluctuates only on the order of minutes to hours. It is entirely non-conceptual, not the result of an agentive, high-level inference, or a judgement about the current behavioural state (whether one is asleep or awake). As such, the experience of bare wakefulness has no grain or internal structure, and can therefore be described as a globalized form of “ultrasmoothness” (see Metzinger, 2003a, sec. 3.2.10, and the references given in footnote 16).

4 Conclusion: MPE as the phenomenology of tonic alertness

This discussion leaves us with a very first, semantically enriched 6D-concept of “minimal phenomenal experience”, which can now, for example, be used as a foundation for systematic psychometric studies. MPE is an inner representation of

---

23How does this relate to the phenomenology of phasic alertness? On a conceptual level, we can interpret tonic alertness as the epistemic capacity to create and sustain sensitivity to incoming perceptual stimuli. This capacity, in turn, is causally related to the epistemic capacity for orientation: the creation of an integrated and egocentric representational space plus the epistemic capacity for executive control of attention. PC1, the phenomenology of wakefulness, is the conscious representation of the first epistemic capacity. Perspectivalness, the phenomenology of being an epistemic agent or a “knowing self”, is the conscious representation of the second and third capacities.

©The author(s). https://philosophymindscience.org ISSN: 2699-0369
tonic alertness, but it comes in variable degrees of constraint satisfaction. It can be coupled with introspective access or not, it can be reflexive or not, and it may be characterized by a variable degree of transparency. Obviously, it can also co-exist with many other forms of conscious content, but because it normally functions as the transparent medium or model of an abstract space in which other contents unfold over time, it is mostly unnoticed. What make it so easy to overlook is the unique combination of transparency with a particularly subtle and abstract kind of phenomenal character, leading to an integration of experiential immediacy with a high degree of counterfactual invariance. In a prototypical form, it would be characterized by a mode-neutral form of bare and functionally autonomous “self-luminous wakefulness” and the accompanying epistemic aspect of subjective confidence, experienced as a quality of “empty cognizance.” The phenomenological prototype would be characterized by a complete absence of time-representation and any form of sensorimotor or high-level cognitive content. Further, there would be an absence of low-level embodiment in the sense of spatiotemporal self-location, interoception, and affective background, as well as of higher levels of selfhood like attentional control and cognitive agency. MPE is non-egoic self-modelling: as such it is atemporal, selfless, and not tied to an individual first-person perspective. It instantiates a unique but highly abstract kind of phenomenal character which can be described on different levels of analysis – for example, as a maximally abstract form of embodiment or bodily self-awareness, as a dynamic representation of tonic alertness, or as an internal model of an epistemic space.

As a representation of a pre-given epistemic space or global medium in which the embodied acquisition of knowledge unfolds over time it has the unique feature of combining the absence of subject/object structure with phenomenal transparency. However, in addition to the specific phenomenological profile already described in sections 2 above and by the resulting 6D-model, it also has certain global aspects that can be introspectively discerned – if there is a prior to initially guide introspective attention, if, as G. E. Moore put it in 1903, “we know that there is something to look for”:

- **Openness:** The phenomenology of tonic alertness is the experience of openness: The organism is now sensitive to incoming stimuli, it can perceive the environment and internal states of its own body, it is open to the world and it represents this very fact. In this sense, MPE is a non-egoic, non-conceptual form of self-knowledge.

- **Simplicity:** All the negative phenomenological characteristics and the Low-Complexity constraint introduced in section 2.1 are satisfied. This

---


©The author(s). https://philosophymindscience.org ISSN: 2699-0369
is the main reason why the phenomenology of tonic alertness is a prime candidate for minimal phenomenal experience.

- **Epistemic capacity:** Third, the phenomenology of tonic alertness is the experience of possessing a set of epistemic capacities. For example, the organism now has the capacity to orientate itself in time and space, or to generate an epistemic first-person perspective by controlling attention (Wiese, 2019). It becomes conscious by non-conceptually knowing about this capacity, but – as a careful investigation into the phenomenology of meditation shows – it does not necessarily exert it. As research on the effortless experience of “pure awareness” in meditation demonstrates, the phenomenology of bare wakefulness is independent of mental agency, time-representation and of self-location in a spatial frame of reference, because it often appears without it. Phenomenologically, the relevant space of possibilities is indivisible, it has no centre and no boundaries, and it is best described as an epistemic space.

- **Expectation of knowledge:** Fourth, there is a specific quality of subjective confidence, an ongoing expectation of epistemic states. Computationally, wakefulness can be described as a statistical hypothesis, a representation of the probability that veridical perception will actually occur (Metzinger, 2019).26

25 On the functional level of analysis, tonic alertness and the orientation reflex can be dissociated: Tonic alertness is what causally enables spatiotemporal self-location and orientation, but it is not identical with it. The conscious experience of wakefulness is a representation of precisely this epistemic capacity, the mere capacity to know where you are, at what point in time you exist, and how you are currently positioned in the order of things. MPE, the phenomenal experience of WAKEFULNESS, is a way in which an organism can non-conceptually know about the existence of this mere capacity, and this is also the reason why the relevant phenomenal character is extremely abstract and subtle. Epistemologically, therefore, wakefulness is an abstract, non-egoic form of self-knowledge. However, this knowledge appears under a specific inner mode of presentation, by using an internal representational format which does not yet involve subject/object structure, time-representation, or spatial embodiment. The ensuing phenomenology can therefore be described as selfless, timeless, and non-spatial, as a model of an epistemic space, a space in which knowledge states can occur.

26 To give an example, it is part of the phenomenology of waking up from dreamless deep sleep that in the very beginning there is a sense of confidence that perceptual states will very soon occur, that not only is one now “open to the world” and knows about one’s epistemic capacity, but that one can and will very soon know where one is, what time it is, and so on. One will also know who one is: Phenomenologically, the gradual transition from unconsciousness to the wake state is characterized by another subtle and intuitive the wake state is characterized by another subtle and intuitive presentiment, something one might perhaps term a phenomenal “foreshadowing of selfhood”. This is an expectation not merely of knowledge, but of egoic self-knowledge. It involves an anticipation of mental agency and the capacity for global self-control (Wiese, 2019), the subjective but as yet non-egoic confidence that a full-blown first-person perspective involving executive control on the mental level plus an extended autobiographical self-model will very soon appear. Therefore, the phenomenal quality of subjective confidence not only relates to knowing


©The author(s). https://philosophymindscience.org ISSN: 2699-0369
- **Spontaneity:** Finally, there is a fifth global aspect characterizing the phenomenology corresponding to the functionally autonomous process of tonic alertness as bare, apparently self-generating and apparently ever-fresh wakeful awareness, and it is related to the absence of time-representation – to the "temporal thinness" mentioned above (Friston et al., 2017). MPE has a specific quality of spontaneity, as it is ahistorical and contains not even an implicit representation of its own causal history. It is the “natural state” – a spontaneous presence of epistemic capacity.

Taken together, these global aspects give us a first more general phenomenological description of the relevant target state: It is a mostly transparent experience of openness to the world, in combination with an abstract, non-conceptual representation of mere epistemic capacity, plus a spontaneously occurring, domain-general sense of confidence, of the likelihood of knowledge – but as yet without object. The target state is experienced as timeless, unconstructed, and functionally autonomous. MPE is aperspectival and it does not instantiate MPS, but it definitely has representational content and a correspondingly unique, phenomenal character *sui generis*. However, the question of whether and in what sense it can count as ”fundamental”, and whether it is the only truly minimal state of consciousness, has not been answered. If the above is correct, phenomenality is a transparent, temporally thin, and global representation of subjective confidence, relative to the abstract possibility of knowledge.

This contribution is the first in a series of publications aiming at advancing our understanding of “minimal phenomenal experience” (MPE). The primary goal of this paper has been to lay the first foundations for a minimal-model explanation of consciousness (cf. section 1.2). One aim is to develop a new entry point for consciousness research by taking a closer look at a specific phenomenological subset, namely “pure consciousness” experiences in meditation (section 2). The first result is that such experiences clearly exist. A second result is that there are actually fewer paradigmatic reports of such experiences in the literature than previously assumed. Third, while such descriptions have strong common phenomenological denominators, they are always “multidimensional” in the sense that they express multiple features and not one single, unequivocal form of phenomenal character. However, the phenomenological prototype to which all such reports

---


©The author(s). https://philosophymindscience.org ISSN: 2699-0369
refer is clearly associated with a fundamental and non-egoic internal representation of tonic alertness (section 3). Its primary dimensions are \textit{Wakefulness} and \textit{Epistemicity}. Possibly, it can conceptually be described as a model of an unpartitioned epistemic space. However, much remains to be clarified. Alongside more detailed philosophical analysis, MPE needs investigating using psychometric, computational, and neuroscientific methods.

**Acknowledgments**

I would like to thank two particularly careful and constructive anonymous reviewers of the Selless Minds project, as well as Raphaël Millière, Adriana Alcaraz Sánchez, Thomas Andrillon, Olivia Carter, Monima Chadha, Tom Clark, Cyril Costines, Nils Donselaar, Martin Dresler, Kristina Eichel, Sascha Fink, Tiziano Furlanetti, Vittorio Gallesse, Jakob Hohwy, Amol Kelkar, Carsten Korth, Rafi Malach, Marcello Massimini, Lucy Mayne, Ulrich Ott, Arnaud Poublan, Danaja Rutar, Katharine Shapcott, Ronald Sladky, Heleen Slagter, Giulio Tononi, Nao Tsuchiya, Wanja Wiese, Jennifer Windt, Mateusz Wozniak and audiences at the 2018 Carnap Lectures in Bochum, the Interdisciplinary College in Günne (IK 2018), the Predictive Processing Group at the Donders Institute in Nijmegen, the Association for the Scientific Study of Consciousness in Kraków (ASSC 22), the Ernst Strüngmann Institute in Frankfurt am Main, the Trends in Mindfulness Conference at the University of Torino, the Max Planck Institute for Human Development in Berlin, the Melbourne Consciousness Group at Monash University, all participants in the MPE workshop in London, Ontario (ASSC 23), and the audience at the Selless Minds workshop at the Frankfurt Institute for Advanced Studies for comments and constructive criticism. I am greatly indebted to Dr Emily Troscianko for editorial support.

**References**


Bachmann, J. K. (2014). \textit{Accounting for pure consciousness: An examination of the ability of the representationalist approach to phenomenal consciousness to account for pure consciousness experiences}. Edmonton: University of Alberta.


©The author(s). https://philosophymindscience.org ISSN: 2699-0369

\footnotesize{[... met the five criteria: adequate reliability and validity, a comprehensive phenomenological assessment, and control for introspective sensitization and demand characteristics” (Pekala, 1991, p. 78).}


©The author(s). https://philosophymindscience.org ISSN: 2699-0369
©The author(s). https://philosophymindscience.org ISSN: 2699-0369


©The author(s). https://philosophymindscience.org ISSN: 2699-0369
Minimal phenomenal experience


©The author(s). https://philosophymindscience.org ISSN: 2699-0369


Open Access
This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

©The author(s). https://philosophymindscience.org ISSN: 2699-0369