



Global Versus Local Theories of Consciousness and the Consciousness Assessment Issue in Brain Organoids

Maxence Gaillard 

Received: 30 June 2023 / Accepted: 18 January 2024
© The Author(s) 2024

Abstract Any attempt at consciousness assessment in organoids requires careful consideration of the theory of consciousness that researchers will rely on when performing this task. In cognitive neuroscience and the clinic, there are tools and theories used to detect and measure consciousness, typically in human beings, but none of them is neither fully consensual nor fit for the biological characteristics of organoids. I discuss the existing attempt relying on the Integrated Information Theory and its models and tools. Then, I revive the distinction between global theories of consciousness and local theories of consciousness as a thought-provoking one for those engaged in the difficult task of adapting models of consciousness to the biological reality of brain organoids. The “micro-consciousness theory” of Semir Zeki is taken as an exploratory path and illustration of a theory defending that minimal networks can support a form of consciousness. I suggest that the skepticism prevailing in the neuroscience community regarding the possibility of organoid consciousness relies on some assumptions related to a globalist account of consciousness

and that other accounts are worth exploring at this stage.

Keywords Brain organoids · Consciousness · Neural correlates of consciousness · Microconsciousness · Integrated information theory

Introduction

Recently, human brain organoids have raised increasing interest from scholars of many fields and a dynamic discussion in bioethics is ongoing. There is a serious concern that these in vitro models of brain development based on innovative methods for three-dimensional stem cell culture might deserve a specific moral status [1, 2]. This would especially be the case if these small stem cell constructs were to develop physiological features of organisms endowed with nervous systems, suggesting that they may be able to feel pain or develop some form of sentience or consciousness. Whether one wants to envision or discard the possibility of conscious brain organoids and whether one wants to acknowledge or dispute its moral relevance, the notion of consciousness is a main pillar of this discussion (even if not the only issue involved [3]). However, consciousness is itself a difficult notion, its nature and definition having been discussed for decades [4, 5]. As a consequence, the ethical debate surrounding brain organoids is deeply entangled with epistemological uncertainty

M. Gaillard (✉)
Centre for Medical Ethics, University of Oslo, Oslo,
Norway
e-mail: m.r.c.gaillard@medisin.uio.no

M. Gaillard
Institut Supérieur de Philosophie, UCLouvain,
Louvain-La-Neuve, Belgium

pertaining to the conceptual underpinnings of the science of consciousness and its empirical endeavor.

It has been argued that neuroethics should circumvent this fundamental uncertainty by adhering to a precautionary principle [6]. Even if we do not know with certainty at which point brain organoids could become conscious, following some experimental design principles would ensure that the research does not raise any ethically problematic features in the years to come. It has also been proposed to redirect the inquiry to the “what-kind” issue (rather than the “whether or not” issue) in order to rely on more graspable features for ethical assessment [7]. These strategies, however, make the epistemological issue even more relevant. The question of whether or not current and future organoids can develop a certain form of consciousness (without presupposing what these different forms of consciousness might be) and how to assess this potentiality in existing biological systems is bound to stay with the field of brain organoid technology for a certain time. Even if it is not for advancing ethical issues, there is a theoretical interest in determining the boundary conditions of consciousness and its potential emergence in artificial entities. Although the methodological and knowledge gap is still wide between the research community on cellular biology and stem cell culture on the one side and the research community on consciousness such as cognitive neuroscience on the other, there will be more and more circulation of ideas and methods in the coming years. The results of this scientific endeavor will, in turn, impact ethics.

In this article, I look back at the history of consciousness research to find new perspectives on this contemporary epistemological conundrum. In particular, I suggest the distinction between “global” theories of consciousness and “local” theories of consciousness as a thought-provoking one for those engaged in the difficult task of adapting models of consciousness to the biological reality of brain organoids. The first section introduces the *consciousness assessment issue* as a general framework and a challenge for any discussion related to the putative consciousness of brain organoids. In the second section, I describe and critically assess the main attempt, so far, at solving the consciousness assessment issue relying on integrated information theory. In the third section, I propose to rely on the distinction between local and global theories of consciousness as a tool to navigate

the theoretical landscape, before turning to the analysis of a notable local theory of consciousness, Semir Zeki’s theory of microconsciousness, in the fourth section. I conclude by drawing the epistemological and ethical lessons from this theoretical exploration.

Detecting consciousness in diverse entities

I delineate here what I call the *consciousness assessment issue*: how to detect the presence or the absence of a possible form of consciousness that could emerge in some brain organoids, assembloids (compounds of organoids), or related technologies. In this sense, consciousness assessment is a scientific problem for which the theoretical and empirical tools are still to be decided. The issue is arising as follows.

- (i). A microphysiological system (e.g., an organoid made of stem cells in culture differentiating and self-organizing in a three-dimensional complex structure) reaches a certain degree of developmental complexity up to the point that it looks like (hence *organ-oid*) a brain, a part of the brain, or a subset of parts of the brain. The very nature of this similarity would deserve more consideration, but let’s say for the moment that the similarity is related to structural or functional aspects of the nervous system at an early stage of development.
- (ii). This phenomenon inclines some scientists and ethicists to envision the possibility of some sort of consciousness or mental activity (sentience, experience...) occurring in microphysiological systems of this kind. If this is ever the case, the most sophisticated systems developed in advanced research settings would be potential candidates. This insight is based on several assumptions, notably the very general one according to which the nervous system plays a key role in consciousness emergence as we know it, but whether it implies a commitment to a purely neurocentric view of consciousness [8, 9] is a point for discussion.
- (iii). As a consequence, the research community wants to assess further this potentiality, for the sake of curiosity and practical, legal reasons – ethical discussions on whether to terminate research or to endow organoids with a specific

moral status might depend on the result of this assessment. For doing that, researchers have to turn to evidence-based methods and empirical measures for describing the actual processes going on and agree on markers that can be identified by observing and manipulating the laboratory entity in consideration.

- (iv). Indeed, in cognitive neuroscience and clinical practice, there are tools and theories built to detect and measure consciousness, typically in human beings. Researchers are already turning to these tools and theories to discuss the potentiality of consciousness emergence [10]. There are at least two difficulties then: none of these tools and theories is fully consensual and none of them is tailored for the new entity of interest (especially in terms of empirical validation and practical measurement, I will discuss these points below).
- (v). The scientific problem of consciousness assessment then states: How to develop a device, akin to a measurement tool with an unambiguous output, that will help us assess whether or not the new entity of concern is developing a form of consciousness?

The consciousness assessment issue in brain organoids is facing two sources of uncertainty. On the one hand, the field of consciousness studies is itself a field of debate with many competing theories. If one is to select a theory of consciousness to assess brain organoids, the number of theories available in the field and the lack of consensus regarding their relevance are certainly confusing [11, 12]. Furthermore, with the exception of some notable efforts [13], there are few signs of advancement towards a possible resolution of the debate which lacks a common nomenclature and fight over the interpretation of the available data. On the other hand, organoids are novel entities and thus pose a specific challenge to scientists studying them.

Analogy is a cognitive strategy commonly employed by scientists when confronted with a novel object and when the methodology to deal with this novel object is not established [14, 15]. Drawing an analogy with a situation or an object with which the research is more familiar, or adapting an established methodology or tool from a related field

of research, is a way to cope with the fundamental uncertainty imposed by novelty.

When mentioning the consciousness assessment issue in cerebral organoids, authors often draw an analogy with the “detection of consciousness” problem in comatose and unresponsive patients [3, 10]. It is indeed not a small achievement of consciousness studies in the past decades to have forced us to revisit the clinical, ontological, and moral status of patients who were supposed to suffer from a complete loss of consciousness before they were assessed again using new tools and models. Thanks to precision gains in imaging tools and refined protocols to obtain functional images of patients’ brains, consciousness researchers have been able to claim that they can, within a certain range of confidence, predict the state of consciousness of patients who are otherwise unable to communicate [16, 17].

This achievement can be seen as a culmination of the branch of neuroimaging research referred to as the “reverse inference” problem, in which researchers are looking at biological signals to infer the mental states of the participants [18]. Although there is a debate on the validity of this kind of inference, there is no absolute rebuttal to the idea that, in principle, reverse inference is possible [19, 20]. The science of consciousness starts from subjective reports to investigate the neural correlates of mental events. Based on participants’ reports on their conscious experience (or on experimental configurations where the conscious experience of the participant is accessible in some way, e.g., a movie shown to the participant [21]), researchers can infer what kind of brain data or pattern of activation is associated with a given state of consciousness. This is of course a complex investigation that requires a subtle delineation of phenomenological concepts. The inference is easier for mapping the motor or perceptual cortex and is growing more and more complex when it comes to disputed psychological notions [22]. Yet at some point – once the science is sufficiently established and correlations are reliable enough – it can be expected that researchers will be able to trust physical measurements in order to make predictions on the conscious state of a system. Reverse inference, in the strict, historical sense, can be seen as a laboratory challenge: there are still conscious participants who can report on their experience to cross-validate the predictions of the model. On the contrary, consciousness detection in comatose

patients raises more epistemological and clinical challenges as a jump into the unknown, because there are no other ways to communicate with patients to validate the tool.

Consciousness detection in organoids faces an additional difficulty, as the biological systems are far from resembling fully grown human brains on which the current models of consciousness in cognitive neuroscience are based. For instance, functional magnetic resonance imaging tools that play a major role in reverse inference research and that have led to some of the most surprising insights into the consciousness of unresponsive patients [17], have been designed for full-scale animal brains and not for non-vascularized tissue in a dish.

The difficulty of assessing consciousness in organoids culminates by combining both sources of uncertainty – the evolving state of the field of consciousness research and the disruptive nature of organoids. Within this perspective, neuroethicists have proposed to rely on several theories of consciousness such as the integrated information theory [3, 10], the global neuronal workspace theory [23, 24], the tempo-spatial theory of consciousness [24], the higher-order theory [7], or the embodied approach [7, 24], among many theories and approaches available. The task is made even more difficult because different theories do not necessarily share the same concepts and definitions, and assessment tools might not even have to rely on one specific theory. However, the work conducted by Lavazza and Massimini [10] and more generally by Lavazza in a series of papers is the only attempt to envision concretely, up to a certain extent, a measurement tool based on one of these theories and it builds from the integrated information theory (IIT).

IIT's ambitions

IIT is a theory of consciousness developed and continuously refined by neuroscientist Tononi and colleagues since the early 2000s [25–27]. In a nutshell, consciousness according to IIT is the ability of a physical system to integrate information. The theory lists the properties that characterize conscious states, so-called “axioms”. These axioms (the exact number and definition of which depends on the version of the theory) state for instance [28] that subjective

experience exists and that a subjective experience is intrinsic (i.e., for a subject of experience), that it is specific (each experience has its own features that make it specific), unitary or integrated (a conscious experience is a unified experience), definite (it is different from other possible experiences), and structured (for instance, a visual experience has generally several features such as shape, color, and motion). Then, the theory identifies under the label “postulates” the corresponding causal properties in the physical substrate instantiating these phenomenal properties. For instance, for a conscious state to be integrated (i.e., one unified experience), each part of the system must be connected with the rest of the system through causal interactions. The level of consciousness is associated with a quantity of information that is integrated in an irreducible manner in a network in which all components have an effect on other components. The theory leans on a mathematical framework interpreting the axioms. This modelling strategy provides a tool to formalize the theory and opens up avenues to empirical applications.

IIT builds from a general theory of conscious experience that could be applied to any physical system. To do so, IIT proponents have proposed a measure of “complexity” to assess the fact that the cognitive system is composed of entangled subsystems that share information, and not independent modules. This indicator is identified with an “index of consciousness,” labeled Φ , that is intended to measure the degree of consciousness of any system of interest, based on the topology of the network of connections in the system. In this context, IIT has been particularly appealing to assess consciousness in brain organoids.

In a landmark article, Lavazza and Massimini [10] hypothesize that Φ can be adapted to assess consciousness potentially emerging in brain organoids. As a proxy for Φ , the authors rely on the “perturbational complexity index” (PCI), which is a measurement of brain complexity proposed by Massimini and colleagues [29]. In line with early versions of IIT [30], brain complexity is here understood as the nature of a system that is both functionally specialized and integrated. In a complex system, one small, local change, will impact the state of the system as a whole. PCI is an attempt to measure complexity in this sense, by stimulating a localized part of the brain and assessing the global impact of this local stimulation. The main tool is transcranial magnetic

stimulation (TMS) combined with electroencephalography recording (EEG): when TMS introduces a local perturbation in the nervous system, this perturbation is likely to lead to massive and unpredictable changes in the system if it is integrated (that would be a sign of consciousness), while it will likely lead to small changes in the global activity pattern if the system is modular (the local perturbation would have only local consequences). The methodology has proven reliable for predicting the capacity for consciousness of awake or anesthetized participants and it has produced interesting results in the clinical assessment of unresponsive patients in a vegetative state [31].

Lavazza & Massimini claim that the PCI method could be adapted to organoids provided that we could replace TMS and EEG with more subtle measurement tools. The fact that PCI has been validated on a specific biological and cognitive system and that its tools (TMS and EEG) at tailored to the human brain is not a definite rebuttal: It would require some work to produce an index that puts all cognitive systems on the same scale and the development of new tools would be a technical challenge but the expectation does not seem unrealistic.

An objection would be that this kind of index and its measurement procedures are as strong as the theory behind it can be. Most scientific instruments are theory-laden (see, e.g., [32]), that is, they are built and validated following the principles of a particular theory, and they can become obsolete with the theory and we know that, according to the regular course of science, all theories have to evolve or become falsified at some point. An index like Φ is very bound to IIT and the fate of the index is in a way committed to the fate of the theory (if one does not endorse IIT, one won't likely endorse Φ). More importantly, the tool is designed to register only what the theory would consider as conscious. As stated above, IIT relies on a certain number of "axioms" of phenomenal experience that determine the conditions under which a physical system is then an eligible candidate for consciousness. When these axioms are updated, does this mean that the phenomena captured are different? Furthermore, the axioms can be challenged [33]. This makes the tool extremely theory-dependent in a dangerous way: if one or several of these axioms do not represent true boundary conditions for conscious states, then the tool would have excluded some conscious states of its potential scope of observation

because of the theoretical assumptions behind its design. For instance, if not all conscious states were composed or structured, a tool designed from the assumption that only a composed state is conscious would miss the detection of some conscious states (e.g., simple forms of subjective experience). This concern is especially relevant in the case of brain organoid consciousness assessment, for we have to be cautious not to discard the most alien and dissimilar forms of consciousness [12, 34].

However, the perturbational complexity index, while inspired by IIT and introduced as a proxy for Φ , might be compatible with other theories such as the global neuronal workspace theory (GNWT) [35]. The history of science has shown that measurement tools can acquire some degree of autonomy once they circulate beyond their theoretical context of emergence (see, e.g., [36]). The reliability of the tool in itself might not be dependent upon the theoretical success of IIT but on its operationalization and validation in the "detection of consciousness" challenge in the clinical context. In the context of a review of consciousness theories, Seth and Bayne state that, "because theories of consciousness are themselves contentious, it seems unlikely that appealing to theory-based considerations could provide the kind of intersubjective validation required for an objective marker of consciousness. Solving the measurement problem thus seems to require a method of validation that is based neither solely on introspection nor on theoretical considerations" [37]. In the absence of a consensus on theories, the research community will likely confront the consciousness assessment issue and especially the challenge of developing a measurement tool by referring to a wide array of relevant concepts and models, even if not always consistent and possibly taken out of their theoretical context. The task would be akin to theoretical and experimental physicists developing a "pidgin" (or creole language) to communicate in order to enable the functioning of big instruments, according to the model of "trading zones" elaborated by historian of science Galison [38]. In the "trading zone" of consciousness research, one will find ambiguous concepts and various experimental protocols and procedures adapted and negotiated between major stakeholders.

When it comes to the consciousness assessment issue in brain organoids, the appeal to tools inspired by IIT is for many good reasons. First, IIT is

currently popular in the field of consciousness studies: it has been selected as one of two main theories in an “adversarial collaboration” to confront their empirical predictions [39]. This current popularity makes IIT an attractive theory today, although popularity should not be taken as a definitive marker of validation or long-lasting authority. Furthermore, the fact that IIT’s ambition is not to be strictly limited to human consciousness and that it wants to be applicable in all physical systems is also a point in favor of its use in unusual contexts such as brain organoids. Then, the measurement index Φ and its counterpart PCI are appealing if the EEG signal could be turned into a reliable “biomarker” of consciousness. Such a marker, if operational for practical measurement, would be a godsend for regulators and bioethicists. According to IIT, Φ could even have the advantage of providing a common measure of consciousness as a natural phenomenon in all kinds of biological and artificial systems, which would mean for instance that we could compare the “consciousness level” of a given brain organoid with the level of consciousness in, say, a fly, a monkey, an X-months old infant, or a locked-in patient (in all these cases, the ethical consequences would be dramatically different). From this, regulators and bioethicists could discuss evidence-based criteria to determine how researchers should behave with the entities of concern: for instance, whether a system for which Φ reaches a certain threshold would deserve to be treated with some respect or on the contrary may be terminated. So many good reasons to adopt a marker that would have all these ideal characteristics (applicable to all kinds of entities and that provides a scale across different levels of consciousness) are also reasons to resist the current proposal and test it against other alternatives. It should also be said that, even if the PCI has made some steps towards clinical validation, IIT does not propose today an empirically validated and untested methodology for the measurement of consciousness in human beings, let alone in other, less familiar systems [40].

In the next section, I will broaden the scope of the reflection by referring to the distinction between global and local theories of consciousness and examine how we might use this distinction as a guide to navigating the “trading zone” of consciousness assessment in brain organoids.

Global versus local theories of consciousness

The distinction between local theories of consciousness and global theories of consciousness is mentioned regularly and more or less formally by actors in the field. For instance, it has been used as a categorization tool in encyclopedia entries presenting a list of theories of consciousness [41]. A recent book by Lau elaborates on this distinction to provide a scale along which different theories of consciousness are distributed [42]. The main idea behind this distinction is that different theories will propose different neural bases for consciousness (or neural correlates of consciousness, NCC). In global theoretical frameworks, the NCC are extended to large parts of the brain, or even the whole brain, while in local theories the NCC are limited to small areas of the brain.

The distinction cannot be quantified and there is no straight line that can be drawn at first sight. There is often no strict delimitation of what “global” means in terms of brain function and the point is not to set a limit to the number of brain areas that should be involved in a network to qualify for local or global. One could ask when an activation starts being global. In a sense, global does not mean that the “whole” brain has to be active for a conscious experience to arise. The nuance is in the contrast: local theories of consciousness look at consciousness as emerging from parts of the nervous system, instead of as the product of a global, widespread pattern. Local theories would put the finger on a specific brain area or a few ones, and consider a strong activity in these areas to be responsible for the emergence of subjective experience. As Lau summarizes: “subjective experiences happen when the right kind of neural activity occurs in the relevant sensory modality... the rest of the brain isn’t really critically involved” [42]. On the contrary, global theories would insist on the idea that the activation should be broadly distributed to enable the emergence of something akin to consciousness. The main concepts put forward by proponents of global theories rely on the distribution of the process: synchronization of activity, long-distance connections, networks of areas that are anatomically distinct, re-entrant loops... All these concepts suggest that there is not only a critical mass of neurons involved but that the key to consciousness lies in the architecture that puts together distinct parts of the nervous system. According to Lau, this idea traces back to

Dennett's "fame in the brain" [4] and the global workspace theory by Baars [43] for which specialized and separated information processing modules broadcast information to a central system. Current versions of the global neuronal workspace make consciousness dependent on the existence of long-range connections between many regions of the brain, including the parietal and prefrontal cortex [44]. Besides, the distinction does not fit with the distinction between "frontal theories" and "parietal theories" (see, e.g., [45]) opposing for instance GNWT for which the activation of the prefrontal cortex is a necessary condition for the emergence of consciousness and IIT insisting on connections inside the visual cortex and related areas. Even a "parietal" or a "frontal" theory will have some global commitments if it attributes consciousness to large activation patterns.

Interestingly, the earliest version of IIT was elaborated from the seminal work of neurobiologist Edelman and its "reentrant dynamic core theory" [25]. According to this approach, consciousness in human beings is dependent on reentry processes made possible by the thalamocortical loop. The information is integrated when it is processed in a network composed of both distributed and interdependent brain regions. The reentry phenomenon is the source of global coordination and synchrony in the brain relying on long-distance connections, and this feature gives rise to a unified experience, or the binding of several elements of perception in one perceptual scene. Several "coalitions of neurons" compete and the successive domination of coalitions explains the variety of conscious experiences. The dynamic core theory insists on the fact that consciousness emerges as the information is processed in the entire thalamocortical network, that is, a global feature of the brain, by contrast with theories that search for *the* locus of consciousness by identifying the brain area responsible for its emergence. Edelman's and Tononi's work converged then on the idea of measuring complexity in a biological system [30]. However, while Edelman's theory relied first on certain neurobiological bases, IIT's approach from axioms to their physical bases suggests that there is no commitment to a particular neuroanatomical realization in the context of IIT. The characterization of IIT as a parietal approach comes from the fact that IIT's proponents have identified the "posterior cortical hot zone" [46] as the complex maximizing Φ . However, the idea according to

which conscious states are integrated and that this integration corresponds to a network of interconnected structures (a "complex") is in itself referring to a kind of globalist account. For instance according to IIT [27], in the brain, the cortex has the kind of physical features that are required for integrating information, while the cerebellum does not because of its modular composition. With regard to PCI, the localized stimulus that leads only to local perturbations is not regarded as a sign of consciousness, while a signature of consciousness would be the observation of massive consequences of a localized stimulus – in other words, the global consequences of local stimulation.

A theory such as Victor Lamme's local recurrency theory [47] would also be interesting because it insists on the temporal dynamics of activation and attributes consciousness to a feedforward wave between the primary visual cortex and temporal areas. While still relatively local compared with GNWT, the concept of recurrence based on connections between different regions refers to more than one specific brain area.

In any case, this global/local distinction has to be taken as a landmark or as a scale rather than a systematic classifier. The distinction is interesting with regard to the issue of organoid consciousness. At first sight, it seems more difficult to build in a dish a system capable of global activation than to replicate the local activity of specific brain regions. Brain organoids are definitely not "mini-brains" in the sense of functional equivalents of full human brains, even at a smaller scale. However, one can relatively easily envision small replicas of brain regions that are realistic enough to exhibit some properties of the regions they model. If we suppose that consciousness emerges when these regions are active, even locally, then we will have to assess this possibility. Of note, in the current state of knowledge, this possibility relies on many unknowns because science still needs to provide a better understanding of the structural and functional correlates of consciousness, not only at the neuronal level but also including the role of the body and the environment. On the contrary, if one trusts global theories only, then one will easily discard the possibility of consciousness emergence in organoids in the years to come. Assembloids (compounds of organoids that replicate distinct brain regions or other organs) might then be a source of concern, but this possibility would still be far remote because the

critical mass of neurons and the long-distance connections that are required for consciousness in biological settings are still out of reach of stem cell biotechnology. Furthermore, if we build our assessment tools (to detect the possible emergence of consciousness in brain organoids) from global theories, then we might miss potentially interesting phenomena that would emerge at a local scale.

In the next section, I will overstate the case on purpose and consider Zeki's theory of minimal consciousness as a clear example of a local theory of consciousness. I do not want to take a stand between global and local theories (nor between IIT and micro-consciousness theory). This article explores instead the meaning of these approaches: how they fit with consciousness assessment in organoids and incline us to look at the problem from a different perspective.

A theory of micro-consciousnesses

The "microconsciousness theory" was proposed by Semir Zeki [48, 49], an expert in functional specialization in the visual brain.¹ Zeki's theory of micro-consciousness stipulates that several consciousnesses can co-exist in the visual system. According to Zeki, consciousness is not a global and unitary phenomenon but involves multiple consciousnesses distributed in distinct processing sites. The visual system is composed of many specialized modules, each exhibiting a sign of partial consciousness, such as consciousness of form, movement, and so on.

The theory starts from the fundamental observation, accumulated over decades and species, that the visual system is composed of rather autonomous subsystems each processing separately information related to color, motion, forms, and so on. In the clinic, dissociations have shown that one subsystem can be impaired while others remain functional. Although contemporary debates are not framed that much in these terms, one of the big challenges for the science of consciousness in the 1990s and early 2000s was the "binding problem" [51]. While I see

a consistent, unified scene (e.g., I see a white cat running from left to right), anatomy and physiology teach us that all these features are processed separately in the visual system (the motion of the cat is encoded in some cortical area, its color in another, etc.) – and the problem makes also sense at a larger scale if we add other sensory modalities. Electrophysiological mapping has shown that distinct neurons are sensitive to orientation, form, color, and that distinct areas are responsible for processing the information related to particular attributes of experience. The binding problem can be formulated in the following way: once all the attributes of the visual world have been delineated in subprocesses that are responsible for, e.g., color/form/movement/location in the brain, how does the nervous system put the pieces together to generate a conscious, unified experience? In this sense, the binding problem equates to the issue of the emergence of consciousness – solving the binding problem would be solving the issue of consciousness. Zeki points to the fact that it is not true by saying that we do not need this kind of binding to have an experience.² Binding is not a necessary condition for the emergence of experience: each subsystem has the ability to generate in parallel a micro-consciousness of its own (e.g., an experience of color, an experience of movement) and one can experience separately color, motion, form. Hence the theory of "microconsciousness."

A motto defended by the microconsciousness theory is that "processing systems are also perceptual systems" [48]. In terms of model building, there is no need to multiply functions. There is no need for a perceptual system on top of the visual system that would turn the visual representation into a percept. If there are some preconscious representations, the percept will be built from this, not from something else.

¹ Zeki is a neurobiologist in London (UCL) who specialized in vision processing [50]. Mainly active during the 1980s-2000s, he is now retired. Most of his expertise was built on electrophysiology of the monkey visual system and he turned in his later years to "neuro-aesthetics" and perception of art.

² Cognitive neuroscientist Dehaene offers also a nice argument for not equating binding and consciousness on different premises [44]. From priming (masking) tasks, we know that consistent representations (i.e., already integrated, unified, where attributes are bounded together) can remain unconscious. For example, chess experts can process the information provided by a subliminal chessboard. That is, we have unconscious representations for which the "binding process" must have already occurred. In other words, binding is neither a necessary (Zeki) nor a sufficient (Dehaene) condition for consciousness.

“We suppose that visual consciousness consists of many, functionally specialized, micro-consciousnesses which are spatially and temporally distributed if they are the result of activity at spatially distributed sites (as in the case of color and motion). This we believe to be the direct consequence of the fact that the several, parallel, multinodal, functionally specialized, and autonomous processing systems are also perceptual ones and that activity at each node of each processing-perceptual system can become perceptually explicit.” [48]

Two main strands of arguments support this view. The first one is taken from psychophysics. Visual perception does not occur as a synchronous phenomenon: some attributes are processed before others, and therefore they can be perceived before others when experimenters find the right manipulations to let that dissonance come to consciousness. “Different processing systems create their corresponding percepts independently and with different delays.” [52]. Especially, participants can perceive location before color, and color before motion (according to the authors, the result is particularly strong for color before motion [52]). Zeki labels this phenomenon the “asynchrony of visual perception.” Over a brief time window, there are several micro-consciousnesses corresponding to different attributes of the visual scene, processed by different subsystems.

The second strand of arguments comes from dissociations on neurological patients. In many neurological syndromes, such as agnosia, patients are unable to perceive a global scene, but they are not deprived of all experience. These patients, while unable to combine their experiences into a whole, are often able to “see and understand what the intact nodes of their processing-perceptual systems allow them to see and understand” [48]. Thanks to the remaining activity in some subsystems of vision processing, they have residual capacities that allow them to see details of a scene that, for them, does not make sense as a whole. This happens when one sees colors or forms without perceiving and identifying familiar objects. For instance, a patient who could not experience shapes and colors because of a lesion of the primary visual cortex could still experience motion – like a person without lesion would perceive, eyes closed, a shadow moving in front of a source of light (according to the

patient’s report). This case is different from the classical interpretation of blindsight as a pathology of consciousness where patients are unable to report a feature of a scene (that is, they are not conscious of it) but still behave in certain experimental conditions as if they could process the information unconsciously [53]. In the case presented by Zeki, the subjective experience is still existing, but reduced to a minimal aspect, as if the only residual processing area after the lesion were also able to produce a minimal feature of consciousness. As a consequence, Zeki suggests that “activity at any given stage of a processing system can have a conscious correlate” [48]. Particular pathways of information processing are responsible for different kinds of subjective experiences. According to Zeki, some patients, affected by specific defects of their visual system, “are capable of a more elementary perceptual experience of the relevant attributes than normals but are nevertheless able to experience something of the relevant attribute” [48], even if their subjective experience does not have the richness of a neurotypical perceptual activity. Such patients are able to “see and experience details of a given attribute without being able to combine the details into a whole and thus experience the whole” [48].

In an overarching article, Zeki introduces a hierarchy between micro-consciousnesses and unified consciousness [49]. When the different attributes coming from different processing subsystems are merged, there is a macro-consciousness, or unified consciousness, at a higher level, that enables the emergence of a global picture. These micro-consciousnesses seem erased when integrated into a macro-consciousness – we have in general the impression of perceiving a moving object, not motion plus an object. However, the author insists on the fact that behind this apparent unity of consciousness, there is disunity, many asynchronous components (micro-consciousness), that are part of the experience. “The quest for the NCC will remain elusive until we acknowledge that consciousness is not a unity, and that there are instead many consciousnesses that are distributed in time and space” [49]. If consciousness is not necessarily unified, because we can be conscious of different aspects at different times, then there are several parts of consciousnesses or snatches of consciousness. In this framework, neither top-down influences nor long-distance connections are required for the emergence of consciousness.

An objection against the micro-consciousness theory is that the processing by subsystems would be a necessary but not sufficient condition for consciousness. A global theorist would hold that “micro-conscious” experiences are not exactly conscious as they become actually conscious only when integrated into a more complex system or broadcasted into a global network. If what is broadcasted is only information about motion, then the subject will be conscious of motion without being conscious of form or color. The debate on sufficient conditions is an empirical question open for debate for which there is scarce, and often disputable, evidence. The micro-consciousness theory is probably not a complete theoretical framework and this article does not want to argue for its validity against other theories of consciousness but more modestly to consider its hermeneutic potential for the consciousness assessment issue in brain organoids.

Discussion

While there is a great interest in neuroethics for human brain organoids and the possibility of these entities deserving a special moral status, a vast majority of actors in the field, especially stem cell researchers and neuroscientists themselves, do not see consciousness of organoids as a realistic possibility and pressing issue. Most of them judge the emergence of high levels of consciousness in artificial entities as very unlikely given the current state of biotechnology and even discard the option as a fantasy [54]. Official reports of academic societies endorse this “nothing to declare” position. For instance, the *International Society for Stem Cell Research* suggests not only that the prospect of conscious in vitro organoids in the foreseeable future is unrealistic, but that it is professional misconduct to communicate publicly on this line: “This is particularly relevant to brain organoids and human-animal chimeras, where any statements implying human cognitive abilities, human consciousness or self-awareness, as well as phrases or graphical representations suggesting human-like cognitive abilities risks misleading the public and sowing doubts about the legitimate nature of such research” [55]. Another recent report states that: “It appears at present that neural organoids have no more moral

standing than other in vitro human neural tissues or cultures. As scientists develop significantly more complex organoids, however, the need to make this distinction will need to be revisited regularly” (although what should count as “significantly more complex” is left to interpretation) [56].

Skeptical accounts (regarding the emergence of human-like consciousness in brain organoids) of this sort are grounded in several assumptions. One of them is a rather anthropocentric or neurotypical concept of consciousness as what matters ethically. The fact that “human-like cognitive abilities” are not in sight does not mean that other, different forms of consciousness do not deserve attention. This is something that the consideration of a broad range of theories of consciousness should encourage us to consider. In general, the field of consciousness studies is full of borderline cases and extreme conditions (from neuropathology, animal experiments, complex experimental design with human subjects) that are very intriguing and should incline us to reexamine our expectations regarding what counts as typical or significant. The position that tells us to postpone the ethical and epistemological reflection while “keeping an eye” on the progress of the technology in very broad terms is problematic in the sense that it does not provide monitoring tools and specific signs of concern.

The skeptical account would somehow follow this line of reasoning: if we want to look for the emergence of a typical, “human-like” form of consciousness in brain organoids, then we have to look for some kind of global activation which of course will not occur because of the limitations of current organoid technology until organoids are composed of multiple realistic interconnected brain systems, like a human brain. The NASEM report [56] states something along this line when it writes that the status of brain organoids is not different from the status of regular in vitro cell culture until “significantly more complex” organoids are grown. It might be an over-interpretation to refer here to IIT’s grounding of consciousness in the “complexity” of a system. One lesson from IIT and its gradual approach might be that a system does not have to be as complex as the human brain to give rise to subjective experience. Less complex systems, but still complex to a certain extent, would have enough “power” to raise concerns – if the

dominant system reaches a certain level of Φ , according to IIT.

This stance would be even stronger with the micro-consciousness theory, according to which we would not have to look for complexity but for the possibility of replicating “perceptual sites” in vitro. In the micro-consciousness theory, experience can emerge from local brain activity. Such a statement would not necessarily be impossible within IIT, although the axioms defining consciousness in this theoretical framework are putting some conditions on what counts as an experience. Depending on how the axioms of IIT are applied, they may impose unnecessary restrictions on the forms of experience that we might want to capture. That would be the case of the axiom of composition posing that all conscious states are structured and composed of several features – in other terms after binding – while microconsciousnesses according to Zeki would occur even at an earlier stage. That would be also the case of the axiom of exclusion, according to which one conscious state “excludes” others so that, if several complexes co-exist in a system, only the one maximizing the value of Φ (labelled Φ^{\max}) will be conscious. In Zeki’s framework, several microconsciousnesses co-exist all along, and it seems that these micro-consciousnesses are interpenetrating and integrated or erased at a higher level when integrated into a macroconscious state.

Raising competing theoretical views on consciousness, even if both views have limitations, has the advantage of questioning the implicit assumptions behind the skeptical account. Human beings typically have two brain hemispheres and see the world as one – and, notably, we are not aware of a boundary between the receptive fields of the primary visual cortex at the border of both hemispheres. The idea that a subsystem responsible for motion detection, color analysis, or shape delineation could give rise to conscious states by themselves is intriguing. What if we try to build organoids that replicate precisely these subsystems? Couldn’t they be subjects of an experience, that we could describe and that could correspond to something that we could experience too? Widening the scope of our models of consciousness will benefit both our discussion of ethical concerns and epistemic curiosity.

We would then have to assess the ethical implications of the different theoretical scenarios. Which exact features of subjective experience would give

rise to ethical consideration and potential moral status? Would a subjective experience staying at a minimal (e.g., perceptual) level be valuable in itself? Sentience, pain, and stress are in general of major concern when it comes to defining the moral status of brain organoids we are interacting with in the laboratory [6, 57]. In this framework, subjective experience is considered morally significant because the experience has a positive or negative value from the viewpoint of the subject of experience. For instance, pain has a negative value that can be detected by the fact that the organism experiencing pain systematically avoids this kind of experience. In other words, the valence of the experience determines its moral significance for a given subject. However, looking for valenced experiences is already conflating a certain number of features of experience (the perceptual content of the experience, the pain or feelings associated with it, and its interest for the organism) that could be analytically distinguished and, potentially, replicated separately in different technological in vitro systems, which we could label “microconscious organoids,” provided the microconsciousness theory is true. In the context of microconscious organoids, one can wonder what would be the moral significance of, for instance, perceptual experience, if it has no valence. “Having an experience of blue” or “being a subject capable of having an experience of blue” is definitely not equivalent to “being a subject of suffering,” but maybe is it still more than being a cell culture in a dish. Even if it were established that a microconscious organoid is sensitive to a certain range of colors, would that be a sufficient condition to impose some restrictions on the use of this organoid for research? Is the status of “subject of experience” something that has to be protected, even if this experience is only perceptual and does not involve pain?

The answers to these questions are not obvious [58] and I cannot explore all their ramifications here. We could however gain from the mobilization of the most local approaches to consciousness, even as a foil, in the following way. If some ethical concern is going to emerge from the potentiality of human brain organoids in the near future, it will not be because of their similarity to a fully-developed, mature human brain, because in vitro models are and will stay very dissimilar to their natural counterparts in many respects [59]. The analogy with various animal nervous systems and developing human brains is

hazardous as well [34]. Thus, starting from the suspicion that even simple systems could acquire a form of microconsciousness whatever the moral significance of this point could be, and then adding the relevant features that would make this experience morally significant is more likely to follow the development of organoid technology. Indeed, a major source of the gain in “complexity” in brain organoid technology relies on the fact that organoids replicating different parts of the brain can be merged in functional assembloids [60, 61]. Even if a biological system is often more than the sum of its parts, a prospective approach with this framework in mind could at least help us identify in advance which assembloids would require a substantive consciousness assessment exercise and which do not.

Acknowledgements I would like to thank the organizers of the research retreat on ELSA of brain organoids in Tübingen for their impulse and continuous efforts. This article was also presented at the “Detecting unusual consciousness” conference at the University of Bonn. I received very insightful comments in both contexts and also from the reviewers of this journal.

Funding Open access funding provided by University of Oslo (incl Oslo University Hospital) The author is employed in EU H2020 HYBRIDA Project, Embedding a comprehensive ethical dimension to organoid based research and related technologies (Grant Agreement 101006012).

Declarations

Conflict of interest The author has no competing interests.No ethics approval was required.Maxence Gaillard is the only author of this work.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Hyun, I., J. Scharf-Deering, and J. Lunshof. 2020. Ethical issues related to brain organoid research. *Brain Research* 1732: 146653. <https://doi.org/10.1016/j.brainres.2020.146653>.
- Koplin, J., O. Carter, and J. Savulescu. 2021. Moral status of brain organoids. In *Rethinking Moral Status*, ed. S. Clarke, H. Zohny, and J. Savulescu, 250–268. Oxford: Oxford University Press.
- Sawai, T., et al. 2021. Mapping the Ethical Issues of Brain Organoid Research and Application. *AJOB Neuroscience* 13 (2): 81–94. <https://doi.org/10.1080/21507740.2021.1896603>.
- Dennett, D. 1993. *Consciousness Explained*. London: Penguin.
- Chalmers, D. 1997. *The Conscious Mind. In Search of a Fundamental Theory*. New York: Oxford University Press.
- Birch, J., and H. Browning. 2021. Neural Organoids and the Precautionary Principle. *AJOB Neuroscience* 21 (1): 56–58. <https://doi.org/10.1080/15265161.2020.1845858>.
- Niikawa, T., Y. Hayashi, J. Shepherd, and T. Sawai. 2022. Human Brain Organoids and Consciousness. *Neuroethics* 15 (5), <https://doi.org/10.1007/s12152-022-09483-1>.
- Noë, A. 2009. *Out of Our Heads: Why You Are not Your Brain, and Other Lessons from the Biology of Consciousness*. New York: Hill and Wang.
- Glannon, W. 2011. *Brain, Body, and Mind: Neuroethics with a Human Face*. New York: Oxford University Press.
- Lavazza, A., and M. Massimini. 2018. Cerebral organoids: Ethical issues and consciousness assessment. *Journal of Medical Ethics* 44 (9): 606–610. <https://doi.org/10.1136/medethics-2017-104555>.
- Francken, J., et al. 2022. An academic survey on theoretical foundations, common assumptions and the current state of consciousness science. *Neuroscience of Consciousness* 2022 (1): niac011. <https://doi.org/10.1093/nc/niac011>.
- Diner, S., and M. Gaillard. 2023. Searching for Consciousness in Unfamiliar Entities: The Need for Both Systematic Investigation and Imagination. 2023. *AJOB Neuroscience* 14 (2): 202–204. <https://doi.org/10.1080/21507740.2023.2188303>.
- Cogitate Consortium. 2023. An adversarial collaboration to critically evaluate theories of consciousness. bioRxiv <https://doi.org/10.1101/2023.06.23.546249>.
- Fox-Keller, E. 2002. *Making Sense of Life: Explaining Biological Development with Models, Metaphors, and Machines*. Cambridge: Harvard University Press.
- Hofstadter, D., and E. Sander. 2013. *Surfaces and Essences: Analogy as the Fuel and Fire of Thinking*. New York: Basic Books.
- Owen, A., M. Coleman, M. Boly, M. Davis, S. Laureys, and J. Pickard. 2006. Detecting Awareness in the Vegetative State. *Science* 313 (5792): 1402–1402. <https://doi.org/10.1126/science.1130197>.
- Laureys, S., and N. Schiff. 2012. Coma and Consciousness: Paradigms (Re)Framed by Neuroimaging. *NeuroImage* 61 (2): 478–491. <https://doi.org/10.1016/j.neuroimage.2011.12.041>.

18. Poldrack, R. 2006. Can Cognitive Processes Be Inferred from Neuroimaging Data? *Trends in Cognitive Science* 10 (2): 59–63. <https://doi.org/10.1016/j.tics.2005.12.004>.
19. Poldrack, R. 2011. Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. *Neuron* 72 (5): 692–697. <https://doi.org/10.1016/j.neuron.2011.11.001>.
20. Machery, E. 2014. In Defense of Reverse Inference. *The British Journal for the Philosophy of Science* 65 (2): 251–267. <https://doi.org/10.1093/bjps/axs044>.
21. Kay, K., and J. Gallant. 2009. I Can See What You See. *Nature Neuroscience* 12 (3): 245. <https://doi.org/10.1038/nrn0309-245>.
22. Figdor, C. 2013. What is the ‘Cognitive’ in Cognitive Neuroscience? *Neuroethics* 6 (1): 105–114. <https://doi.org/10.1007/s12152-012-9157-5>.
23. Lavazza, A. 2021. Consciousnessoids: clues and insights from human cerebral organoids for the study of consciousness. *Neuroscience of Consciousness* 7 (2), <https://doi.org/10.1093/nc/niab029>.
24. Zilio, F., and A. Lavazza. 2023. Consciousness in a Rotor? Science and Ethics of Potentially Conscious Human Cerebral Organoids. *AJOB Neuroscience*. 14 (2): 178–196. <https://doi.org/10.1080/21507740.2023.2173329>.
25. Edelman, G., and G. Tononi. 2000. *A Universe of Consciousness: How Matter Becomes Imagination*. New York, NY: Basic Books.
26. Tononi G. 2004. An information integration theory of consciousness. *BMC Neuroscience* 5 (42), <https://doi.org/10.1186/1471-2202-5-42>.
27. Tononi, G., M. Boly, M. Massimini, and C. Koch. 2016. Integrated information theory: From consciousness to its physical substrate. *Nature Review Neuroscience* 17 (7): 7. <https://doi.org/10.1038/nrn.2016.44>.
28. Albantakis, L. et al. 2022. Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms. *arXiv.org*, <https://arxiv.org/abs/2212.14787v1>. Accessed 20.01.2024.
29. Massimini, M., M. Boly, A. Casali, M. Rosanova, and G. Tononi. 2009. A perturbational approach for evaluating the brain’s capacity for consciousness. *Progress in Brain Research* 177: 201–214. [https://doi.org/10.1016/S0079-6123\(09\)17714-2](https://doi.org/10.1016/S0079-6123(09)17714-2).
30. Tononi, G., and G. Edelman. 1998. *Consciousness and Complexity Science* 282 (5395): 1846–1851. <https://doi.org/10.1126/science.282.5395.1846>.
31. Casali, A., et al. 2013. A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine* 5 (198): 198ra105. <https://doi.org/10.1126/scitranslmed.3006294>.
32. Chalmers, A. 2003. The Theory-Dependence of the Use of Instruments in Science. *Philosophy of Science* 70 (3): 493–509. <https://doi.org/10.1086/376924>.
33. Bayne, T. 2018. On the axiomatic foundations of the integrated information theory of consciousness. *Neuroscience of Consciousness* 1: niy007. <https://doi.org/10.1093/nc/niy007>.
34. Diner, S. 2023. Potential Consciousness of Human Cerebral Organoids: On Similarity-Based Views in Precautionary Discourse. *Neuroethics* 16 (3): 23. <https://doi.org/10.1007/s12152-023-09533-2>.
35. Farisco, M., and J.-P. Changeux. 2023. About the compatibility between the perturbational complexity index and the global neuronal workspace theory of consciousness. *Neuroscience of Consciousness* 1: niad016. <https://doi.org/10.1093/nc/niad016>.
36. Bickle, J. 2016. Revolutions in Neuroscience: Tool Development. *Frontiers in Systems Neuroscience* 10: 24. <https://doi.org/10.3389/fnsys.2016.00024>.
37. Seth, A., and T. Bayne. 2022. Theories of consciousness. *Nature Reviews Neuroscience* 23 (7): 439–452. <https://doi.org/10.1038/s41583-022-00587-4>.
38. Galison, P. 1997. *Image and Logic: A Material Culture of Microphysics*. Chicago: Chicago University Press.
39. Melloni, L., et al. 2023. An adversarial collaboration protocol for testing contrasting predictions of global neuronal workspace and integrated information theory. *PLoS ONE* 18 (2): e0268577. <https://doi.org/10.1371/journal.pone.0268577>.
40. Cerullo, M. 2015. The Problem with Phi: A Critique of Integrated Information Theory. *PLOS Computational Biology* 11 (9): e1004286. <https://doi.org/10.1371/journal.pcbi.1004286>.
41. Kouider, S. 2009. Neurobiological Theories of Consciousness. In *Encyclopedia of Consciousness*, ed. W. Banks, 87–100. Elsevier.
42. Lau, H. 2022. *In Consciousness we Trust: The Cognitive Neuroscience of Subjective Experience*. Oxford, New York: Oxford University Press.
43. Baars, B. 1988. *A cognitive theory of consciousness*. Cambridge: Cambridge University Press.
44. Dehaene, S. 2014. *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*. New York, NY: Penguin Books.
45. Northoff, G., and V. Lamme. 2020. Neural signs and mechanisms of consciousness: Is there a potential convergence of theories of consciousness in sight? *Neuroscience and Biobehavioral Reviews* 118: 568–587. <https://doi.org/10.1016/j.neubiorev.2020.07.019>.
46. Koch, C., M. Massimini, M. Boly, and G. Tononi. 2016. Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience* 17 (5): 5. <https://doi.org/10.1038/nrn.2016.22>.
47. Lamme, V. 2006. Towards a true neural stance on consciousness. *Trends in Cognitive Science* 10 (11): 494–501. <https://doi.org/10.1016/j.tics.2006.09.001>.
48. Zeki, S., and A. Bartels. 1999. Toward a theory of visual consciousness. *Consciousness and Cognition* 8 (2): 225–259. <https://doi.org/10.1006/ccog.1999.0390>.
49. Zeki, S. 2003. The duality of consciousness. *Trends in Cognitive Science* 7 (5): 214–218. [https://doi.org/10.1016/s1364-6613\(03\)00081-0](https://doi.org/10.1016/s1364-6613(03)00081-0).
50. Zeki, S., J. Watson, C. Lueck, K. Friston, C. Kennard, and R. Frackowiak. 1991. A Direct Demonstration of Functional Specialization in Human Visual Cortex. *Journal of Neuroscience* 11 (3): 641–649. <https://doi.org/10.1523/JNEUROSCI.11-03-00641.1991>.
51. Treisman, A. 1996. The binding problem. *Current Opinion in Neurobiology* 6 (2): 171–178. [https://doi.org/10.1016/s0959-4388\(96\)80070-5](https://doi.org/10.1016/s0959-4388(96)80070-5).
52. Zeki, S., and A. Bartels. 1998. The asynchrony of consciousness. *Proceedings: Biological Sciences* 265

- (1405): 1583–1585. <https://doi.org/10.1098/rspb.1998.0475>.
53. Derrien, D., C. Garric, C. Sergent, and S. Chokron. 2022. The nature of blindsight: implications for current theories of consciousness. *Neuroscience of Consciousness* 1: niab043. <https://doi.org/10.1093/nc/niab043>.
54. Lavazza, A., and A. Chinaia. 2023. Human cerebral organoids: The ethical stance of scientists. *Stem Cell Research & Therapy* 14 (1): 59. <https://doi.org/10.1186/s13287-023-03291-x>.
55. ISSCR. 2021. International Society for Stem Cell Research Guidelines for Stem Cell Research and Clinical Translation. Available at <https://www.isscr.org/guidelines>. Accessed 20.01.2024.
56. National Academies of Sciences, Engineering, and Medicine. 2021. *The Emerging Field of Human Neural Organoids, Transplants, and Chimeras: Science, Ethics, and Governance*. Washington, DC: The National Academies Press, <https://doi.org/10.17226/26078>.
57. Koplin, J., and J. Savulescu. 2019. Moral Limits of Brain Organoid Research. *Journal of Law and Medical Ethics* 47 (4): 760–767. <https://doi.org/10.1177/1073110519897789>.
58. Shepherd, J. 2018. *Consciousness and Moral Status*. London: Routledge. <https://doi.org/10.4324/9781315396347>.
59. Gaillard, M., and M. Botbol-Baum. 2022. Pursuit of Perfection? On Brain Organoids as Models. *AJOB Neuroscience* 13 (2): 79–80. <https://doi.org/10.1080/21507740.2022.2048735>.
60. Paşca, S. 2018. The rise of three-dimensional human brain cultures. *Nature* 553 (7689): 437–445. <https://doi.org/10.1038/nature25032>.
61. Miura, Y., M.-Y. Li, O. Revah, S.-J. Yoon, G. Narazaki, and S. Paşca. 2022. Engineering brain assembloids to interrogate human neural circuits. *Nature Protocols* 17 (1): 15–35. <https://doi.org/10.1038/s41596-021-00632-z>.

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