A Synergistic Workspace for Human Consciousness Revealed by Integrated Information Decomposition

Andrea I. Luppi[1,2]*, Pedro A.M. Mediano[3], Fernando E. Rosas[4,5,6], Judith Allanson[1,7], John D. Pickard[1,8,9], Robin L. Carhart-Harris[4,10], Guy B. Williams[1,8], Michael M Craig[1,2], Paola Finoia[1], Adrian M. Owen[11], Lorina Naci[12], David K. Menon[2,8], Daniel Bor[3], & Emmanuel A. Stamatakis[2]

[1]Department of Clinical Neurosciences, University of Cambridge, Cambridge, CB2 0QQ, United Kingdom.
[2]University Division of Anaesthesia, School of Clinical Medicine, University of Cambridge, Cambridge, CB2 0QQ, United Kingdom.
[3]Department of Psychology, University of Cambridge, Cambridge, CB2 3EB, United Kingdom.
[4]Center for Psychedelic Research, Department of Brain Science, Imperial College London, London W12 0NN, United Kingdom.
[5]Center for Complexity Science, Imperial College London, London SW7 2AZ, United Kingdom.
[6]Data Science Institute, Imperial College London, London SW7 2AZ, United Kingdom.
[7]Department of Neurosciences, Cambridge University Hospitals NHS Foundation, Addenbrooke's Hospital, Hills Rd, CB2 0SP, Cambridge, United Kingdom
[8]Wolfson Brain Imaging Centre, University of Cambridge, Cambridge, CB2 0QQ, United Kingdom.
[9]Division of Neurosurgery, School of Clinical Medicine, University of Cambridge, Addenbrooke's Hospital, Hills Rd, CB2 0SP, Cambridge, United Kingdom
[10]Psychedelics Division - Neuroscape, Department of Neurology, University of California, San Francisco, CA 94143
[11]Department of Psychology and Department of Physiology and Pharmacology, The Brain and Mind Institute, N6A 5B7 University of Western Ontario, London, Ontario, Canada.
[12]Trinity College Institute of Neuroscience, School of Psychology, Lloyd Building, Trinity College Dublin, Dublin 2, Ireland.
*Correspondence: al857@cam.ac.uk

Abstract

A central goal of neuroscience is to understand how the brain orchestrates information from multiple input streams into a unified conscious experience. Here, we address two fundamental questions: how is the human information-processing architecture functionally organised, and how does its organisation support consciousness? We combine network science and a rigorous information-theoretic notion of synergy to delineate a “synergistic global workspace”, comprising gateway regions that gather synergistic information from specialised modules across the brain. This information is then integrated within the workspace and widely distributed via broadcaster regions. Through functional MRI analysis, we show that gateway regions of the synergistic workspace correspond to the brain’s default mode network, whereas broadcasters coincide with the executive control network. Demonstrating the empirical relevance of our proposed architecture for neural information processing, we show that loss of consciousness due to general anaesthesia or disorders of consciousness corresponds to a diminished ability of the synergistic workspace to integrate information, which is restored upon recovery. Thus, loss of consciousness coincides with a breakdown of information integration within the synergistic workspace of the human brain. This work contributes to conceptual and empirical reconciliation between two prominent scientific theories of consciousness, the Global Neuronal Workspace and Integrated Information Theory. Taken together, this work provides a new perspective on the role of prominent resting-state networks within the human information-processing architecture, while also advancing our understanding of how the human brain supports consciousness through the synergistic integration of information.

Keywords: synergy, global workspace, integrated information, brain networks, anaesthesia, disorders of consciousness

Introduction

Humans and other vertebrates rely on a centralised nervous system to process information from the environment, obtained from a wide array of sensory sources. Information from different sensory sources must eventually be combined - and integrated - with the organism's memories and goals, in order to guide adaptive behaviour effectively [1]. However, understanding how the brain's information-processing architecture enables the integration of information remains a key open challenge in neuroscience [2,3]. Theoretical and empirical work in cognitive neuroscience indicates that information processed in parallel by domain-specific sensory modules needs to be integrated within a multimodal "central executive" [4]. Indeed, recent work has identified subsets of regions that are consistently recruited across a variety of tasks [5–7], situated at the convergence of multiple anatomical, functional, and neurochemical hierarchies in the brain [8–20].

Prominent theories in cognitive and computational neuroscience have also proposed that global integration of information from diverse sources plays a fundamental role in relation to human consciousness [21,22]. The influential Global Neuronal Workspace Theory (GNWT) focuses on the process by which specific neural information becomes available for conscious access, as occurring through the global integration induced by a "global workspace" [23–26]. Within the workspace, relevant information from different sources is integrated and subsequently broadcasted back to the entire brain, in order to inform further processing and achieve "experiential integration" of distributed cortical modules into a coherent whole [24,25,27]. Thus, the global workspace is attributed both the role of integrator, and the role of orchestrator of cognitive function. Also highlighting the importance of integration, the prominent Integrated Information Theory (IIT) [21,28,29] posits that the degree of consciousness in a system is determined by its "integrated information": the amount of intrinsic information generated by the dynamics of the system considered as a whole, over and above the information generated by the dynamics of its individual constituent parts [21,28–30]. Thus, this notion of integrated information corresponds to the extent to which "the whole is greater than the sum of its parts" [31].

Therefore, leading theoretical accounts of consciousness converge on this point: consciousness critically depends on the capability for global integration across a network of differentiated modules. Despite agreeing on the fundamental importance of information integration [32], these theories differ on its specific role and corresponding neural mechanisms. In contrast to GNWT's account, whereby integration is viewed as a necessary - but not sufficient - prerequisite step on the way to broadcasting and consciousness, IIT proposes a more fundamental identity between consciousness and the integration of information, but without specifying a formal architecture for this process: that is, according to IIT any system that integrates information will thereby be conscious, regardless of its specific organisation [31]. Seen under this light, it becomes apparent that IIT and GNWT are actually addressing different aspects of consciousness, and their views of integration are different but potentially complementary.

Crucially, our ability to make sense of any information-processing architecture is limited by our understanding of the information that is being processed. An elegant formal account of information in distributed systems - such as the human brain - is provided by the framework of Partial Information Decomposition (PID) [33] which extends the formalism of Shannon mutual information by demonstrating that not all information is equal. Mutual information quantifies the reduction in uncertainty about one variable, when another variable is taken into account. In the case when more than one source of information is present, PID demonstrates that two sources can possess information about a given target that is unique (each source provides independent information), redundant (the same information is provided by both sources) or synergistic (complementary information, a higher-order kind of information that is available only when both sources are considered together). As an

example, humans have two sources of visual information about the world: two eyes. The information that is lost when one eye is closed is called the "unique information" of that source - information that cannot be obtained from the remaining eye. The information that one still has when one eye is closed is called "redundant information" - because it is information that is carried equally by both sources. This provides robustness: you can still see even after losing one eye. However, losing one eye also deprives you of stereoscopic information about depth. This information does not come from either eye alone: you need both, in order to perceive the third dimension. Therefore, this is called the "synergistic information" between the sources - the extra advantage that is derived from combining them. Synergistic information therefore reflects the meaning of integration-as-cooperation, whereby elements are distinct from each other, but complementary [34].

Adding to the rich literature that addresses neural information from the perspective of encoding and decoding of task variables [35], there is growing appreciation that distinct types of information – as identified by information decomposition - may play a key role in the distributed information-processing architecture of the brain [34,36–42,42,43]. Information decomposition can be applied to neural data from different scales, from electrophysiology to functional MRI, with or without reference to behaviour [34]. When behavioural data are taken into account, information decomposition can shed light on the processing of "extrinsic" information, understood as the translation of sensory signals into behavioural choices across neurons or regions [41,43–45]. However, information decomposition can also be applied to investigate the "intrinsic" information that is present in the brain's spontaneous dynamics in the absence of any tasks, in the same vein as resting-state "functional connectivity" and methods from statistical causal inference such as Granger causality [46]. In this context, information processing should be understood in terms of the dynamics of information: where and how information is stored, transferred, and modified [34]. Specifically, since the future state of the brain is at least in part determined by its previous state, it is possible to view the future state of neural units (be they regions or neurons) as the target, and ask how it is determined by the same units' previous state, and the previous state of other units, which become the sources of information. Then, redundancy between two units occurs when their future spontaneous evolution is predicted equally well by the past of either unit. Synergy instead occurs when considering the two units together increases the mutual information between the units' past and their future – suggesting that the future of each is shaped by its interactions with the other. At the microscale (e.g., for spiking neurons) this phenomenon has been suggested as reflecting "information modification" [36,40,47]. Synergy can also be viewed as reflecting the joint contribution of parts of the system to the whole, that is not driven by common input[48].

By applying a recent generalisation of PID for timeseries data – known as Integrated Information Decomposition [42,49] – we developed an information-resolved approach to decompose the information carried by brain dynamics and their intrinsic fluctuations [50]. Traditional measures of statistical association ("functional connectivity") cannot disentangle synergy and redundancy; in fact, recent work has demonstrated that functional connectivity predominantly reflects redundant interactions [34,50,51]. In contrast, applying our information-resolved framework to functional MRI recordings of the human brain revealed that different regions of the human brain predominantly rely on different kinds of information for their interactions with other regions. Through this approach, we identified a "synergistic core" of brain regions supporting higher-level cognitive functions in the human brain through the synergistic integration of information [50]. Similar results of a synergistic architecture were recently and independently obtained using a different decomposition (based on entropy rather than mutual information) [51].

We also observed that a synergy-based measure of emergent dynamics in functional MRI recordings is disrupted in patients suffering from chronic disorders of consciousness [52]. Building on these findings, it is natural to ask whether this synergistic core could correspond

to the brain's global workspace. Furthermore, given that the views on information integration put forward by GNWT and IIT are potentially complementary, an important challenge to move the field forward is to leverage both accounts into a unified architecture that could explain empirical effects observed in neuroimaging data.

Therefore, this work sets out to address two fundamental questions of contemporary neuroscience:

1. How is the cognitive architecture of the human brain functionally organised, from an information-theoretic standpoint? Specifically, what brain regions does it involve, and what are the roles of the two kinds of information integration proposed by GNWT and IIT within this architecture?
2. How are different types of information in the brain related to human consciousness?

To address these questions, and provide an information-resolved view of human consciousness, here we study three resting-state fMRI datasets: (i) N=100 subjects from the Human Connectome Project; (ii) N=15 healthy volunteers who were scanned before and after general anaesthesia with the intravenous propofol as well as during post-anaesthetic recovery [53]; (iii) N=22 patients suffering from chronic disorders of consciousness (DOC) as a result of severe brain injury [53]. By comparing functional brain scans from transient anaesthetic-induced unconsciousness and from the persistent unconsciousness of DOC patients, which arises from brain injury, we can search for common brain changes associated with loss of consciousness – thereby disambiguating what is specific to loss of consciousness.

Results

Adopting an information-resolved view, we propose to divide the information-processing stream within the human brain in three key stages: (i) gathering of information from multiple distinct modules into a workspace; (ii) integration of the gathered information within the workspace; and (iii) global information broadcasting to the rest of the brain. Furthermore, we propose that while all workspace regions are involved in stage (ii), they are differentially involved in stages (i) and (iii).

The existence of a synergistic workspace and these three processing stages can be seen as emerging from a trade-off between performance and robustness that is inherent to distributed systems. Theoretical work in cognitive science [26] and the field of distributed signal processing [54,55] has long recognised the computational benefits of combining multiple distinct processing streams. However, having a single source of inputs to and outputs from the workspace introduces what is known as a "single point of failure," which can lead to catastrophic failure in case of damage or malfunction [56]. Therefore, a natural solution is to have not a single but multiple units dedicated to gathering and broadcasting information, respectively, thereby forming a workspace that can be in charge of synthesising the results of peripheral processing [57].

Pertaining to Stage (ii), we previously identified which regions of the human brain predominantly entertain synergistic interactions, and thus are most reliant on combining information from other brain regions [50] (Figure S1). The key signature of workspace regions is to have a high prevalence of synergistic (compared to redundant) functional interactions, and therefore the synergy-rich regions that we discovered are ideally poised as GNW candidates. Here, we consider the architecture of the global workspace more broadly, and combine Integrated Information Decomposition with graph-theoretical principles to bring

insights about processing stages (i) and (iii) (Figure 1). We term this proposal the "Synergy-Φ-Redundancy" neurocognitive architecture (SAPHIRE) (Figure 1).

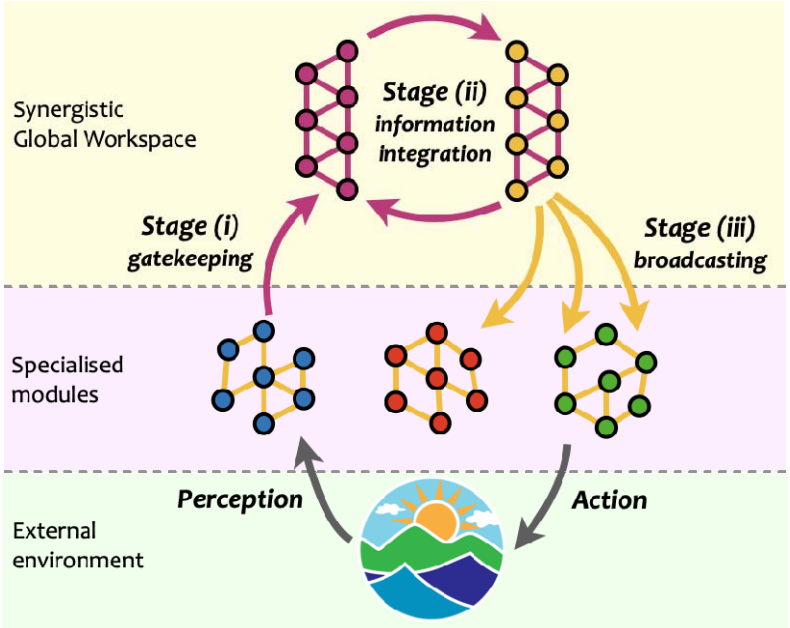

Figure 1. Schematic of the proposed SAPHIRE neurocognitive architecture. Below, specialised modules characterised by robust redundant functional interactions process information about the environment. Information is then collected by workspace gateways through synergistic interactions [Stage (i)]; synergistic interactions integrate information within the synergistic global workspace [Stage (ii)]; workspace broadcasters spread the integrated information back to the specialised modules, through redundant interactions [Stage (iii)], for further processing and to guide behaviour. Orange links represent redundant interactions, and violet links represent synergistic interactions. Grey arrows represent interactions between the system and its environment.

We note that brain regions through which information gains access to the workspace should exhibit synergistic functional interactions that are widely distributed across the brain, as - by definition - the workspace gathers and synthesises information from a multiplicity of diverse brain modules. Thus, we postulate that regions that mediate the access to the synergistic workspace are functionally connected with multiple modules within networks of synergistic interactions, synthesising incoming inputs from diverse sources [58,59]. We refer to such regions as *gateways* (Figure 1, violet nodes). In contrast, the process of broadcasting information corresponds to disseminating multiple copies of the same information from the workspace to many functionally adjacent brain regions. Therefore, broadcaster regions also have functional interactions with many different modules, but of non-synergistic, redundant interactions: "redundancy" accounts for the fact that multiple copies of the same information are being distributed. These regions are designated as *broadcasters* (Figure 1, orange nodes).

One approach to operationalise these ideas is by leveraging well-established graph-theoretical tools. Here, we propose to assess the diversity of intermodular functional connections using the *participation coefficient* [60] which captures to what extent a given node connects to many modules beyond its own (Materials and Methods). Note that this is different from the node *strength,* which captures a region's total amount of connectivity, and which we used to identify which regions belong to the synergistic workspace (see Materials

and Methods and Ref. [50]); the participation coefficient instead quantifies the diversity of modules that a region is connected to. Therefore, gateways are identified from rs-fMRI data as brain regions that (a) belong to the workspace (i.e., have high total synergy), and (b) have a highly-ranked participation coefficient in terms of synergistic functional interactions. Conversely, broadcasters are global workspace regions (i.e., also having high synergy) that have a highly-ranked participation coefficient rank for redundant interactions.

In other words, we identify the synergistic workspace as regions where synergy predominates, which as our previous research has shown, are also involved with high-level cognitive functions and anatomically coincide with transmodal association cortices at the confluence of multiple information streams [50]. This is what we should expect of a global workspace. Subsequently, to discern broadcasters from gateways within the synergistic workspace, we seek to encapsulate the meaning of a "broadcaster" in information terms. We argue that this corresponds with making the same information available to multiple modules. Sameness of information corresponds to redundancy, and connection with multiple modules can be reflected in the network-theoretic notion of participation coefficient. Thus, a broadcaster is a region in the synergistic workspace (i.e., a region with strong synergistic interactions) that in addition has a high participation coefficient for its redundant interactions.

To explore these hypotheses, we quantified synergistic and redundant interactions between 454 cortical and subcortical brain regions [61,62] based on resting-state functional MRI data from 100 subjects of the Human Connectome Project [50]. Specifically, we systematically applied Integrated Information Decomposition to groups of four variables: the past and future of region X, and the past and future of region Y, for all combinations of X and Y. This provided us with a full decomposition of how information is jointly conveyed by X and Y (redundantly, uniquely, or synergistically) across time. In particular, following our previous work [50] we focused on the persistent synergy (henceforth simply synergy) and persistent redundancy (henceforth simply redundancy), which correspond to the information that is always carried synergistically (respectively, redundantly) by X and Y.

We then subdivided the brain into the well-established resting-state networks identified by Yeo and colleagues [63] plus an additional subcortical module [64]. Based on this partition into modules, we identified gateways and broadcasters by comparing the participation coefficients of synergistic versus redundant interactions, for brain regions belonging to the synergistic workspace previously identified (we show a significant correlation for participation coefficient obtained from modules defined a priori as the well-known resting-state networks, or defined in a data-driven fashion from Louvain community detection [65] (Figure S2).

Intriguingly, our results reveal that gateways reside primarily in the brain's default mode network (Figure 2B, violet). In contrast, broadcasters are mainly located in the executive control network, especially lateral prefrontal cortex (Figure 2B, orange). Remarkably, the latter results are in line with Global Neuronal Workspace Theory, which consistently identifies lateral prefrontal cortex as a major broadcaster of information [24,66].

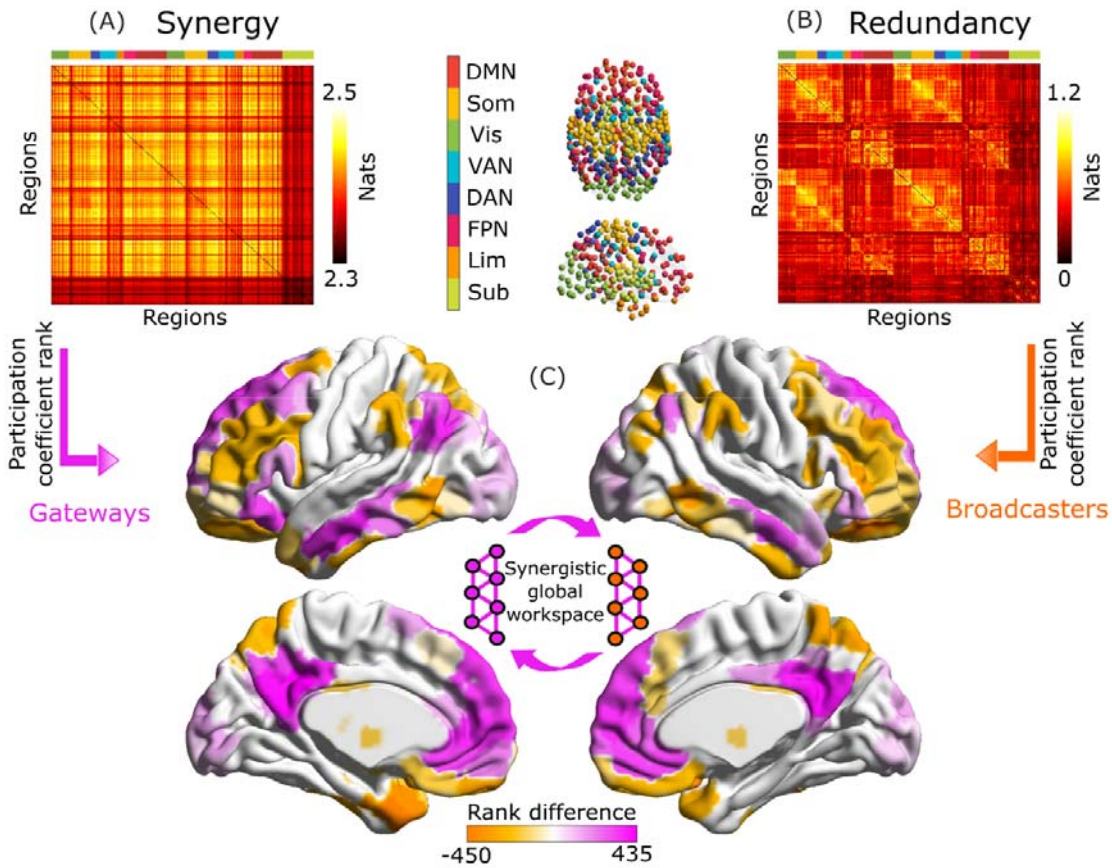

Figure 2. Gateways and broadcaster regions identified by their network connectivity profiles. (A) Group-average matrix of synergistic interactions between regions of the 454-ROI augmented Schaefer atlas. (B) Group-average matrix of redundant interactions. For ease of visualization, the colorscale in (B) pertains to log-transformed values. We highlighted modular allegiance to the canonical resting-state networks by using the colour scheme shown in between A and B. (C) Regions are identified as gateways (violet) or broadcasters (orange) based on the difference between rank of participation coefficient for synergy and redundancy, (only shown for brain regions identified as belonging to the synergistic global workspace, as per Luppi et al., 2022). Violet indicates synergy rank > redundancy rank, corresponding to workspace regions that combine information of many brain modules (gateways); orange indicates the opposite, identifying workspace regions that broadcast information to many modules. Inset: illustration of the synergistic workspace. Legend: DMN, default mode network. Som, somatomotor network. Vis, visual network. VAN, ventral attention network. DAN, dorsal attention network. FPN, fronto-parietal control network. Lim, limbic network. Sub, subcortical network (comprised of 54 regions of the Tian 2020 atlas [64]). These results were also replicated using an alternative parcellation with 232 cortical and subcortical nodes (Figure S3).

Information decomposition identifies a synergistic core supporting human consciousness

Having introduced a taxonomy within the synergistic global workspace based on the distinct informational roles of different brain regions, we then sought to investigate their role in supporting human consciousness. Given the importance attributed to integration of information by both GNWT and IIT, we expected to observe reductions in integrated information within the areas of the synergistic workspace associated with loss of consciousness. Furthermore, we also reasoned that any brain regions that are specifically involved in supporting consciousness should "track" the presence of consciousness: the reductions should occur regardless of how loss of consciousness came about, and they should be restored when consciousness is regained.

We tested these hypotheses with resting-state fMRI from 15 healthy volunteers who were scanned before, during, and after anaesthesia with the intravenous agent propofol, as well as 22 patients with chronic disorders of consciousness (DOC) [53]. Resting-state fMRI data were parcellated into 400 cortical regions from the Schaefer atlas, and 54 subcortical brain regions from the Tian atlas (same parcellation as for the previous analysis). Building on the IIT literature, which provides a formal definition of integrated information, we assessed integration corresponding to conscious activity via two alternative metrics: the well-known whole-minus-sum Φ measure introduced by Balduzzi and Tononi [31], and the "revised Φ" (Φ_R) measure recently introduced by Mediano, Rosas and colleagues [49] (Materials and Methods and Figure 3). Being demonstrably non-negative, this revised measure overcomes a major conceptual limitation of the original formulation of integrated information [49].

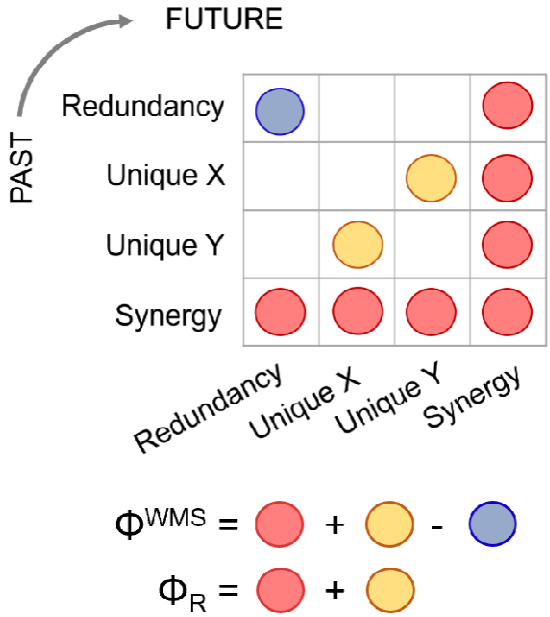

Figure 3. Integrated information decomposition. Integrated information decomposition identifies how two sources X and Y jointly convey information across time, corresponding to all possible 4x4 combinations of redundancy, unique information (of X and of Y), and synergistic information. This decomposition highlights why

the original whole-minus-sum Φ measure introduced by Balduzzi and Tononi can be negative: because it involves the subtraction of the persistent redundancy that is present in the system, leading to negative values in systems that are redundancy-dominated. This shortcoming can be corrected with the revised measure of Φ, termed Φ_R.

For each subject, we computed the integrated information between each pair of BOLD signal timeseries, resulting in a 454-by-454 matrix of integrated information between brain regions. Treating this matrix as an (undirected) network enabled us to study consciousness-related changes in integrated information across conditions, which were analysed using the Network Based Statistic correction for multiple comparisons [67]. Importantly, since we are interested in changes that are shared between the DOC and propofol datasets, we computed edge-level statistics using a composite null hypothesis test designed to detect such shared effects (Materials and Methods).

Analysis based on Φ_R revealed a widespread reorganisation of integrated information throughout the brain when comparing awake volunteers against DOC patients, with both increases and decreases being observed ($p < 0.001$; Figure 4A). Likewise, propofol anaesthesia was also characterised by significant changes in integrated information between brain regions, both when compared with pre-anaesthetic wakefulness ($p < 0.001$; Figure 4B) and post-anaesthetic recovery ($p < 0.001$; Figure 4C).

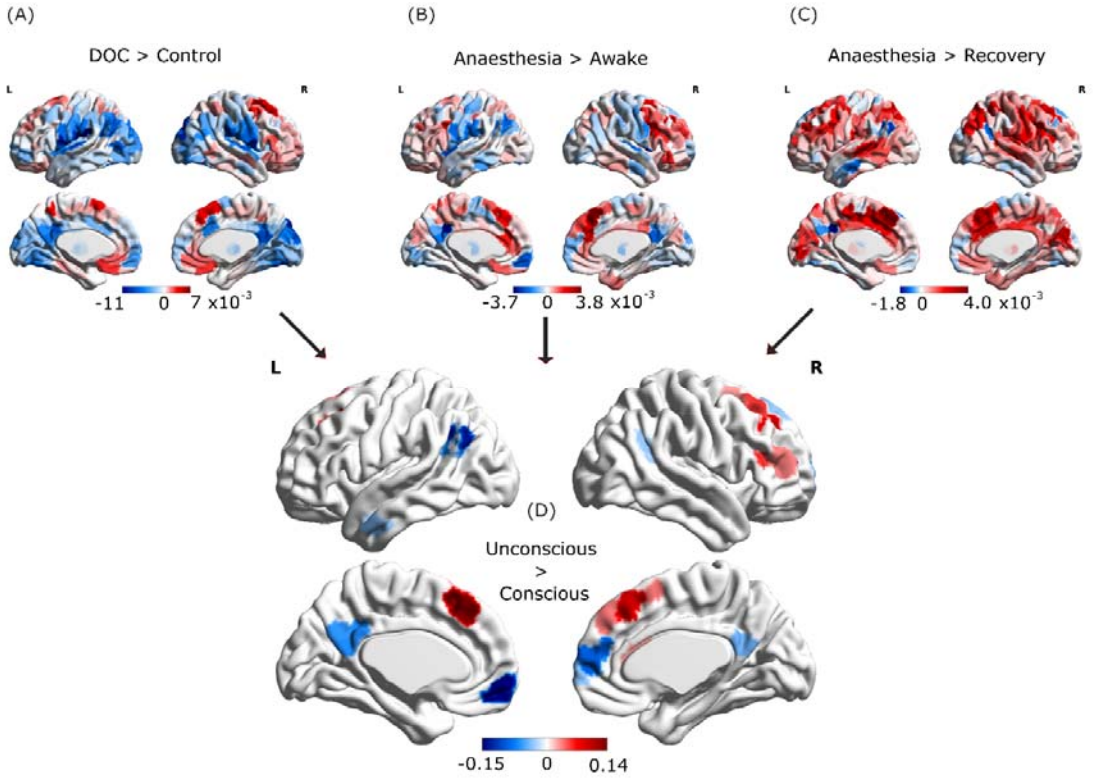

Figure 4. Loss of consciousness induces similar reorganisation of cortical integrated information across anaesthesia and disorders of consciousness. Top: Brain regions exhibiting overall NBS-corrected increases (red) and decreases (blue) in integrated information exchange when consciousness is lost. (A) DOC patients minus awake healthy volunteers; (B), propofol anaesthesia minus pre-induction wakefulness; (C) propofol-anaesthesia minus post-anaesthetic recovery. (D) Overlaps between the three contrasts in (A-C), showing increases and decreases that are common across anaesthesia and disorders of consciousness.

Our analysis identified a number of the Φ_R connections that were reduced when consciousness was lost due to both anaesthesia and brain injury, and were restored during post-anaesthetic recovery - as we had hypothesised (Figure 4D). Remarkably, almost all regions showing consistent decreases in Φ_R when consciousness was lost were members of the global synergistic workspace, and specifically located in the default mode network (bilateral precuneus and medial prefrontal cortex) - and bilateral inferior parietal cortex – although left temporal cortices were also involved (Figure 4D). Additionally, some connections exhibited increases in Φ_R during loss of consciousness, and were restored upon recovery (Figure 4D), including areas in frontal cortex - especially lateral prefrontal cortex. Nevertheless, the overall balance was in favour of reduced integrated information: sum of F-scores associated with significant edges = -25.37 (Figure S4).

These results were in contrast with the analysis based on the original formulation of Φ introduced by Balduzzi and Tononi [31], which did not identify any reductions in integrated information that were common across anaesthesia and disorders of consciousness, instead only identifying common increases (Figure S5A).

Having identified the subset of brain regions that are reliably associated with supporting human consciousness in terms of their integrated information, the last step of our analysis was to leverage the architecture proposed above to understand their role in our information-based view of the global workspace. Since IIT predicts that loss of consciousness corresponds to reductions in integrated information, we focused on regions exhibiting reliable reductions in Φ_R when consciousness is lost (whether due to anaesthesia or DOC), which were restored upon recovery (shown in blue in Figure 4D).

Remarkably, our whole-brain results show that Φ_R disconnections induced by loss of consciousness play the role of gateway nodes (Figure 5A, violet) rather than broadcaster nodes (Figure 5A, orange) according to our previous identification of gateways and broadcasters from the Human Connectome Project dataset (see Figure 2B, violet regions). Indeed, all reductions occur specifically within the default mode network (Figure 5B). Thus, these results suggest that loss of consciousness across anaesthesia and disorders of consciousness would correspond to anterior-posterior disconnection - in terms of integrated information - between DMN nodes that act as gateways into the synergistic workspace.

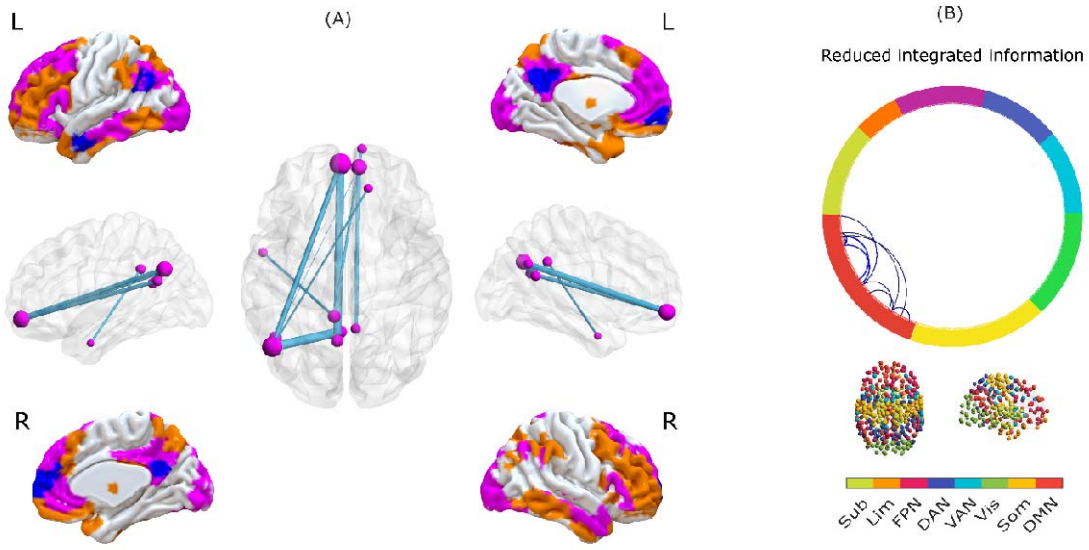

Figure 5. Synergistic core of human consciousness. (A) Surface representations of medial and lateral views of the brain (L indicates left, R indicates right). Colours indicate brain regions that belong to the synergistic workspace, as identified from HCP data. Orange indicates broadcasters, and violet indicates gateways. Blue indicates regions that exhibit an overall significant reduction in integrated information across anaesthesia and disorders of consciousness. All blue regions overlap with violet ones. The network representation indicates edges with significantly reduced integrated information (Φ_R) during both propofol anaesthesia and disorders of consciousness. The color of the nodes (violet or orange) indicates whether the corresponding regions are workspace gateways (violet) or broadcasters (orange); all regions are gateways (violet). (B) Circular graph representation of significant reductions in integrated information (Φ_R) between brain regions, observed in all three contrasts, indicating membership of canonical resting-state networks. Connections indicate pairs of regions with a significant decrease of integrated information. Colour of the circle border indicates the RSN affiliation of the corresponding regions. Legend: DMN, default mode network. Som, somatomotor network. Vis, visual network. VAN, ventral attention network. DAN, dorsal attention network. FPN, fronto-parietal control network. Lim, limbic network. Sub, subcortical network (comprised of 54 regions of the Tian 2020 atlas).

Robustness and sensitivity analysis

To ensure the robustness of our results to analytic choices, we also replicated them using an alternative cortical parcellation of lower dimensionality: we used the Schaefer scale-200 cortical parcellation [62] complemented with the scale-32 subcortical ROIs from the Tian subcortical atlas [64] (Figure S5B). Additionally, we also show that our results are not dependent on the choice of parameters in the NBS analysis, and are replicated using an alternative threshold definition for the connected component (extent rather than intensity) or a more stringent value for the cluster threshold ($F > 12$) (Figure S5C-D). Importantly, whereas the increases in Φ_R are not the same across different analytic approaches, reductions of Φ_R in medial prefrontal and posterior cingulate/precuneus are reliably observed, attesting to their robustness.

Discussion

Architecture of the synergistic global workspace

This paper proposes an informational perspective on the brain's functional architecture at the macroscale, which leverages insights from network science and a refined understanding of neural information exchange. The synergy-Φ-redundancy (SAPHIRE) architecture posits the

existence of a “synergistic workspace” of brain regions characterised by highly synergistic global interactions, which we previously showed to be composed by prefrontal and parietal cortices that are critical for higher cognitive functions [50]. This workspace is further functionally decomposed by distinguishing gateways, which bring information from localised modules into the workspace, and broadcasters, which disseminate multiple copies of workspace information back to low-level regions.

Remarkably, our results on the HCP dataset show that the proposed operationalisation of gateways and broadcasters corresponds to the distinction between the brain’s default mode network and executive control network, respectively. This data-driven identification of workspace gateways and broadcasters with the DMN and FPN provides a new framework to explain well-known functional differences between DMN and FPN, based on their distinct and complementary roles within the brain’s synergistic global workspace, which is discussed below.

The fronto-parietal executive control network (FPN) mainly comprises lateral prefrontal and parietal cortices, and it is associated with performance of a variety of complex, cognitively demanding tasks [68–70]. A key component of this network is lateral prefrontal cortex (LPFC). Based on theoretical and empirical evidence, as summarised in a recent review of GNWT [24], this region is posited to play a major role in the global workspace, as a global broadcaster of information. Remarkably, this is precisely the role that our results assigned to LPFC, based on its combined information-theoretic and network properties. These results are also consistent with recent insights from network neuroscience, which indicate that the FPN is ideally poised to steer whole-brain dynamics through novel trajectories, in response to complex task demands [70,71]. Specifically, by broadcasting to the rest of the brain information that has been integrated within the workspace, the FPN may act as global coordinator of subsequent whole-brain dynamics.

On the other hand, the default mode network comprises posterior cingulate and precuneus, medial prefrontal cortex, and inferior parietal cortices [72–74]. This network, whose constituent regions have undergone substantial developments in the course of human evolution [75,76], was found to occupy a crucial position at the convergence of functional gradients of macroscale cortical organization [15,77,78], forming a structural and functional core of the human brain [79–81], in line with its recently observed involvement in cognitive tasks [82–84]. In particular, the DMN is prominently involved in self-referential processing [85,86], and ‘mental-time-travel’ [87] or episodic memory and future-oriented cognition [88–91]. Its posterior regions in particular, act as relays between the neocortex and the hippocampal memory system [89]. Thus, in terms of both neuroanatomical connectivity and functional engagement, the DMN is uniquely positioned to integrate and contextualise information coming into the synergistic global workspace (e.g. from sensory streams) by combining it with rich information pertaining to one’s past experiences and high-level mental models about ‘self’ and world [78,92–96] - coinciding with the results of the present analysis, which identify DMN nodes as gateways of inputs to the synergistic global workspace.

It is worth noting that the role of the FPN-DMN tandem in supporting consciousness has been suggested by Shanahan’s hypothesis of a ‘connective core’ along the brain’s medial axis [97]. While Shanahan’s hypotheses were primarily based on structure, in this work we combine novel information-theoretic tools to confirm and expand the connective core hypothesis from a functional, information-centric perspective, in a way that differentiates the multiple roles played by the different regions that together comprise this connective core (Figures 2,5).

Integrated Information Decomposition of human consciousness

After identifying the neuroanatomical-functional mapping of the synergistic workspace in terms of gateways and broadcasters, we sought to identify their role in supporting human consciousness. Considering integrated information as a marker of consciousness (without necessarily assuming the two to be identical), we focused on identifying regions where information integration is reduced when consciousness is lost (regardless of its cause, be it propofol anaesthesia or severe brain injury), and restored upon its recovery. Our results indicate that brain regions exhibiting consciousness-specific reductions in integrated information coincide with major nodes of the synergistic global workspace.

Intriguingly, we found that the main disruptions of information integration were localised in gateway nodes, rather than broadcasters. Thus, loss of consciousness in both anaesthesia and disorders of consciousness could be understood as a breakdown of the entry points to the "synergistic core" (Figure 5), which becomes unable to properly integrate inputs for the workspace. Importantly, the original "whole-minus-sum" Φ introduced by Balduzzi and Tononi [31] did not show consistent reductions during loss of consciousness. Thus, the present results demonstrate the empirical validity of the "revised" measure, Φ_R, in addition to its theoretical soundness [49]. Since workspace gateway regions coincide with the brain's default mode network, these results are also in line with recent evidence that information content and integrative capacity of the DMN are compromised during loss of consciousness induced by both anaesthesia and severe brain injury [53,98–106], and even COVID-19 [107]. Due to its prominent role in self-referential processing [86], breakdown of DMN connectivity within the synergistic workspace may be seen as a failure to integrate one's self-narrative into the "stream of consciousness", in the words of William James.

This notion is further supported by focusing on reductions of integrated information during anaesthesia compared with wakefulness. In addition to the synergistic core, overall reductions are also observed in a set of thalamic, auditory and somatomotor regions, largely resembling the brain regions that stop responding to sensory (auditory and noxious) stimuli once the brain reaches propofol-induced saturation of EEG slow-wave activity (SWAS [108]). Although there was no EEG data available to confirm this, the doses of propofol employed in the present study are compatible with the doses of propofol at which SWAS has been shown to arise [109], and therefore it is plausible that our participants also reached SWAS and the loss of brain responsiveness it indicates. Thus, both resting-state integration of information between brain regions, as well as stimulus-evoked responses within each region [108], converge to indicate that propofol disrupts further processing of thalamocortical sensory information – a phenomenon termed "thalamocortical isolation" [108]. We propose that as the thalamus and sensory cortices lose their ability to respond to stimuli, they cease to provide information to the synergistic core of the global workspace, resulting in a disconnection from the external world and presumably loss of consciousness.

These results testify to the power of the Integrated Information Decomposition framework: by identifying the information-theoretic components of integrated information, we have been able to obtain insights about human consciousness that remained elusive with alternative formulations, and could not be captured via standard functional connectivity or related methods. Thus, our findings are consistent with the notion that the global workspace is relevant for supporting consciousness in the human brain, in line with the proposal that *"[...] unconsciousness is not necessarily a complete suppression of information processing but rather a network dysfunction that could create inhospitable conditions for global information exchange and broadcasting"* [24]. GNWT postulates a key role for the global workspace in supporting consciousness: consistent with this theory, we find that several nodes of the synergistic global workspace become disconnected from each other in terms of integrated information when consciousness is lost, especially between anterior and posterior regions (Figure 4, brain networks). Thus, these are brain regions that (i) belong to the synergistic global workspace; (ii) exhibit overall reductions of integrated information when

consciousness is lost; and (iii) are disconnected from other regions of the synergistic workspace when consciousness is lost. The brain regions satisfying these three conditions therefore meet the criteria for constituting an interconnected "synergistic core" of workspace regions supporting human consciousness.

Limitations and future directions

in order to obtain high spatial resolution for our identification of workspace regions, here we relied on the BOLD signal from functional MRI, which is an indirect proxy of underlying neuronal activity, with limited temporal resolution. However, we sought to alleviate potential confounds by deconvolving the hemodynamic response function from our data with a dedicated toolbox [110] (Materials and Methods), which has been previously applied both in the context of information decomposition [50], as well as anaesthetic-induced loss of consciousness [111], and disorders of consciousness [52]. Additionally, the present results of an overall Φ_R reduction are also broadly in line with those of a previous study [112], whose measure of synergy-redundancy balance showed, in ECoG recordings of non-human primates, a broadband shift away from synergy during anaesthesia.

It is also worth bearing in mind that our measure of integrated information between pairs of regions does not amount to measuring the integrated information of the brain as a whole, as formally specified in the context of Integrated Information Theory [31] - although we do show that the average integrated information between pairs of regions is overall reduced across the whole brain. We also note that our revised measure of integrated information is based on IIT 2.0 [31], due to its computational tractability; as a result, it relies on a conceptually distinct understanding of integrated information from the more recent IIT 3.0 [113] and IIT 4.0 [114] versions. Thus, these limitations should be borne in mind when seeking to interpret the present results in the context of IIT. Indeed, future work may benefit from seeking convergence with recent advances in the characterization of emergence, which is related to integrated information [115–118].

Likewise, it is not our intention to claim that Φ_R and synergy, as measured at the level of regional BOLD signals, represent a direct cause of consciousness, or are identical to it. Rather, our work is intended to use these measures similarly to the use of sample entropy and Lempel-Ziv complexity for BOLD signals [53,119]: as theoretically grounded macroscale indicators, whose empirical relationship to consciousness may point towards the relevant underlying neural phenomena. In other words, while our results do show that BOLD-derived Φ_R tracks the loss and recovery of consciousness, we do not claim that they are the cause of it: only that an empirical relationship exists, which is in line with what we might expect on theoretical grounds. Future work will be required to identify whether this empirical relationship also holds at the microscale, and whether the causal mechanisms that induce loss of consciousness are also causally responsible for loss of integrated information.

Intriguingly, although we have focused on anaesthetic-induced decreases in integrated information, due to IIT's prediction that this is what should occur during loss of consciousness, our results also indicate concomitant increases of integrated information – possibly reflecting compensatory attempts, although we refrain from further speculation (Figure 4). Interestingly, increases appear to coincide with broadcaster nodes of the synergistic workspace. In particular, even though lateral prefrontal cortices are among the regions most closely associated with the global neuronal workspace in the literature [24,66], our results indicate a paradoxical net increase in lateral prefrontal integrated information during anaesthesia and DOC. We interpret this qualitatively different behaviour as indicating that different subsets of the global workspace may be differentially involved in supporting consciousness.

However, we note that, whereas the decreases in integrated information were robust to the use of different analytic approaches (e.g., use of a different parcellation or different NBS

threshold), the increases that we observed were less robust, with no region consistently showing increases in integrated information (Figure S5B-D). Nevertheless, both this phenomenon and the meaning of increased integrated information between brain regions deserve further investigation. Indeed, dreaming during anaesthesia has been reported to occur in up to 27% of cases [120], and behaviourally unresponsive participants have been shown to perform mental imagery tasks during anaesthesia, both of which constitute cases of disconnected consciousness [121]. Thus, although our doses of propofol were consistent with the presence of SWAS, we cannot exclude that some of our participants may have been merely disconnected but still conscious, possibly driving the increases we observed.

More broadly, future research may also benefit from characterising the role of the synergistic workspace in the states of altered consciousness induced e.g. by psychedelics [122–124], especially since prominent involvement of the DMN has already been identified [125,126]. Likewise, the use of paradigms different from resting-state, such as measuring the brain's spontaneous responses to engaging stimuli (e.g. suspenseful narratives [127] or engaging movies [128]) may provide evidence for a more comprehensive understanding of brain changes during unconsciousness. Likewise, it will be of great interest to investigate whether and how reorganization of the synergistic global workspace is reflected in other indicators of consciousness [129], such as the brain's response to external perturbations – such as the EEG response to brief magnetic pulses used to compute the Perturbational Complexity Index, one of the most discriminative indices of consciousness available to date [130–134].

The PCI is used as a means of assessing the brain's current state, but stimulation protocols can also be adopted to directly induce transitions between states of consciousness. In rodents, carbachol administration to frontal cortex awakens rats from sevoflurane anaesthesia[135], and optogenetic stimulation was used to identify a role of central thalamus neurons in controlling transitions between states of responsiveness[136,137]. Additionally, several studies in non-human primates have now shown that electrical stimulation of the central thalamus can reliably induce awakening from anaesthesia, accompanied by the reversal of electrophysiological and fMRI markers of anaesthesia [138–143]. Finally, in human patients suffering from disorders of consciousness, stimulation of intra-laminar central thalamic nuclei was reported to induce behavioural improvement [144], and ultrasonic stimulation [145,146] and deep-brain stimulation are among potential therapies being considered for DOC patients [147,148]. It will be of considerable interest to determine whether our corrected measure of integrated information and topography of the synergistic workspace also restored by these causal interventions.

Additionally, the reliance here on 'resting-state' data without external stimuli may have resulted in an overestimation of the DMN's role in consciousness, and an under-estimation of the FPN (including lateral PFC), given their known different recruitment during no-task conditions [72]. Indeed, recent efforts have been carried out to obtain a data-driven characterisation of the brain's global workspace based on regions' involvement across multiple different tasks [5]. This work is complementary to ours in two aspects: first, the focus of Deco *et al.* [5] is on the role of the workspace related to cognition, whereas here we focus primarily on consciousness. Second, by using transfer entropy [149,150] as a measure of functional connectivity, Deco and colleagues [5] assessed the directionality of information exchange – whereas our measure of integrated information is undirected, but are able to distinguish between different kinds of information being exchanged and integrated. Thus, different ways of defining and characterising a global workspace in the human brain are possible, and can provide complementary insights about distinct aspects of the human neurocognitive architecture. Indeed, transfer entropy can itself be decomposed into information-dynamic atoms through Partial Information Decomposition and Integrated Information Decomposition [33,34,49,151]; ΦID can further decompose the Normalised Directed Transfer Entropy measure used by Deco *et al* [5], as recently demonstrated [152]. We look forward to a more refined conceptualization of the synergistic workspace architecture that

takes into account both information types and the directionality of information flow – especially in datasets with higher temporal resolution.

Looking forward, growing evidence indicates an important role for brain dynamics and time-resolved brain states in supporting cognition [7,18,153–160] and consciousness [53,106,161–168]. Therefore, time-resolved extensions of our framework, such as developed by Varley and colleagues [42], may shed further light on the dynamics of the synergistic workspace, especially if combined with neuroimaging modalities offering higher temporal resolution, such as magneto- or electroencephalography. More broadly, a key strength of our proposed cognitive architecture is its generality: being entirely grounded in the combination of information theory and network science, it could be applied to shed light on cognition in humans and other organisms [169], but also to inspire further development of artificial cognitive systems [34,170–174].

Conclusion

Overall, we have shown that powerful insights about human consciousness and neurocognitive architecture can be obtained through the information-resolved approach, afforded by the framework of Integrated Information Decomposition. Importantly, the proposed criteria to identify gateways, broadcasters, and the synergistic workspace itself, are based on practical network and information-theoretic tools, which are applicable to a broad range of neuroimaging datasets and neuroscientific questions. These findings bring us closer to a unified theoretical understanding of consciousness and its neuronal underpinnings - how mind arises from matter.

Materials and Methods

The propofol and DOC patient functional data employed in this study have been published before [53,119,122,175–178]. For clarity and consistency of reporting, where applicable we use the same wording as our previous work [53,122,175].

Anaesthesia Data: Recruitment

The propofol data were collected between May and November 2014 at the Robarts Research Institute in London, Ontario (Canada) [53]. The study received ethical approval from the Health Sciences Research Ethics Board and Psychology Research Ethics Board of Western University (Ontario, Canada). Healthy volunteers (n=19) were recruited (18–40 years; 13 males). Volunteers were right-handed, native English speakers, and had no history of neurological disorders. In accordance with relevant ethical guidelines, each volunteer provided written informed consent, and received monetary compensation for their time. Due to equipment malfunction or physiological impediments to anaesthesia in the scanner, data from n=3 participants (1 male) were excluded from analyses, leaving a total n=16 for analysis [53,122,175].

Anaesthesia Data: Procedure

Resting-state fMRI data were acquired at different propofol levels: no sedation (Awake), and Deep anaesthesia (corresponding to Ramsay score of 5). As previously reported [53,122,175], for each condition fMRI acquisition began after two anaesthesiologists and one anaesthesia nurse independently assessed Ramsay level in the scanning room. The anaesthesiologists and the anaesthesia nurse could not be blinded to experimental condition, since part of their role involved determining the participants' level of anaesthesia. Note that the Ramsay score is designed for critical care patients, and therefore participants did not receive a score during

the Awake condition before propofol administration: rather, they were required to be fully awake, alert and communicating appropriately. To provide a further, independent evaluation of participants' level of responsiveness, they were asked to perform two tasks: a test of verbal memory recall, and a computer-based auditory target-detection task. Wakefulness was also monitored using an infrared camera placed inside the scanner.

Propofol was administered intravenously using an AS50 auto syringe infusion pump (Baxter Healthcare, Singapore); an effect-site/plasma steering algorithm combined with the computer-controlled infusion pump was used to achieve step-wise sedation increments, followed by manual adjustments as required to reach the desired target concentrations of propofol according to the TIVA Trainer (European Society for Intravenous Aneaesthesia, eurosiva.eu) pharmacokinetic simulation program. This software also specified the blood concentrations of propofol, following the Marsh 3-compartment model, which were used as targets for the pharmacokinetic model providing target-controlled infusion. After an initial propofol target effect-site concentration of 0.6 μg mL^{-1}, concentration was gradually increased by increments of 0.3 μg mL^{1}, and Ramsay score was assessed after each increment: a further increment occurred if the Ramsay score was lower than 5. The mean estimated effect-site and plasma propofol concentrations were kept stable by the pharmacokinetic model delivered via the TIVA Trainer infusion pump. Ramsay level 5 was achieved when participants stopped responding to verbal commands, were unable to engage in conversation, and were rousable only to physical stimulation. Once both anaesthesiologists and the anaesthesia nurse all agreed that Ramsay sedation level 5 had been reached, and participants stopped responding to both tasks, data acquisition was initiated. The mean estimated effect-site propofol concentration was 2.48 (1.82- 3.14) μg mL^{-1}, and the mean estimated plasma propofol concentration was 2.68 (1.92- 3.44) μg mL^{-1}. Mean total mass of propofol administered was 486.58 (373.30- 599.86) mg. These values of variability are typical for the pharmacokinetics and pharmacodynamics of propofol. Oxygen was titrated to maintain SpO2 above 96%.

At Ramsay 5 level, participants remained capable of spontaneous cardiovascular function and ventilation. However, the sedation procedure did not take place in a hospital setting; therefore, intubation during scanning could not be used to ensure airway security during scanning. Consequently, although two anaesthesiologists closely monitored each participant, scanner time was minimised to ensure return to normal breathing following deep sedation. No state changes or movement were noted during the deep sedation scanning for any of the participants included in the study [53,122,175]. Propofol was discontinued following the deep anaesthesia scan, and participants reached level 2 of the Ramsey scale approximately 11 minutes afterwards, as indicated by clear and rapid responses to verbal commands. This corresponds to the "recovery" period [176].

Anaesthesia Data: Design

As previously reported [53,122,175], once in the scanner participants were instructed to relax with closed eyes, without falling asleep. Resting-state functional MRI in the absence of any tasks was acquired for 8 minutes for each participant, in each condition. A further scan was also acquired during auditory presentation of a plot-driven story through headphones (5-minute long). Participants were instructed to listen while keeping their eyes closed. The present analysis focuses on the resting-state data only; the story scan data have been published separately, and will not be discussed further here.

Anaesthesia Data: FMRI Data Acquisition

As previously reported [53,122,175], MRI scanning was performed using a 3-Tesla Siemens Tim Trio scanner (32-channel coil), and 256 functional volumes (echo-planar images, EPI) were collected from each participant, with the following parameters: slices = 33, with 25% inter-slice gap; resolution = 3mm isotropic; TR = 2000ms; TE = 30ms; flip angle = 75 degrees; matrix size = 64x64. The order of acquisition was interleaved, bottom-up. Anatomical scanning was also performed, acquiring a high-resolution T1- weighted volume (32-channel

coil, 1mm isotropic voxel size) with a 3D MPRAGE sequence, using the following parameters: TA = 5min, TE = 4.25ms, 240x256 matrix size, 9 degrees flip angle [53,122,175].

Disorders of Consciousness Patient Data: Recruitment

A total of 71 DOC patients were recruited from specialised long-term care centres from January 2010 to December 2015 [53,122,175]. Ethical approval for this study was provided by the National Research Ethics Service (National Health Service, UK; LREC reference 99/391). Patients were eligible to be recruited in the study if they had a diagnosis of chronic disorder of consciousness, provided that written informed consent to participation was provided by their legal representative, and provided that the patients could be transported to Addenbrooke's Hospital (Cambridge, UK). The exclusion criteria included any medical condition that made it unsafe for the patient to participate, according to clinical personnel blinded to the specific aims of the study; or any reason that made a patient unsuitable to enter the MRI scanner environment (e.g. non-MRI-safe implants). Patients were also excluded based on substantial pre-existing mental health problems, or insufficient fluency in the English language prior to their injury. After admission to Addenbrooke's Hospital, each patient underwent clinical and neuroimaging testing, spending a total of five days in the hospital (including arrival and departure days). Neuroimaging scanning took place at the Wolfson Brain Imaging Centre (Addenbrooke's Hospital, Cambridge, UK), and medication prescribed to each patient was maintained during scanning.

For each day of admission, Coma Recovery Scale-Revised (CRS-R) assessments were recorded at least daily. Patients whose behavioural responses were not indicative of awareness at any time, were classified as UWS. In contrast, patients were classified as being in a minimally conscious state (MCS) if they provided behavioural evidence of simple automatic motor reactions (e.g., scratching, pulling the bed sheet), visual fixation and pursuit, or localisation to noxious stimulation. Since this study focused on whole-brain properties, coverage of most of the brain was required, and we followed the same criteria as in our previous studies [53,122,175]; before analysis took place, patients were systematically excluded if an expert neuroanatomist blinded to diagnosis judged that they displayed excessive focal brain damage (over one third of one hemisphere), or if brain damage led to suboptimal segmentation and normalisation, or due to excessive head motion in the MRI scanner (exceeding 3mm translation or 3 degrees rotation). Of the initial sample of 71 patients who had been recruited, a total of 22 adults (14 males; 17–70 years; mean time post injury: 13 months) meeting diagnostic criteria for Unresponsive Wakefulness Syndrome/Vegetative State or Minimally Conscious State due to brain injury were included in this study. In addition to the researcher and radiographer, a research nurse was also present during scanning. Since the patients' status as DOC patients was evident, no researcher blinding was possible.

Disorders of Consciousness Patient Data: FMRI Data Acquisition

As previously reported [53,122,175], resting-state fMRI was acquired for 10 minutes (300 volumes, TR=2000ms) using a Siemens Trio 3T scanner (Erlangen, Germany). Functional images (32 slices) were acquired using an echo planar sequence, with the following parameters: 3 x 3 x 3.75mm resolution, TR = 2000ms, TE = 30ms, 78 degrees FA. Anatomical scanning was also performed, acquiring high-resolution T1-weighted images with an MPRAGE sequence, using the following parameters: TR = 2300ms, TE = 2.47ms, 150 slices, resolution 1 x 1 x 1mm.

Functional MRI preprocessing and denoising

The functional imaging data were preprocessed using a standard pipeline, implemented within the SPM12-based (http://www.fil.ion.ucl.ac.uk/spm) toolbox CONN

(http://www.nitrc.org/projects/conn), version 17f [179]. The pipeline comprised the following steps: removal of the first five scans, to allow magnetisation to reach steady state; functional realignment and motion correction; slice-timing correction to account for differences in time of acquisition between slices; identification of outlier scans for subsequent regression by means of the quality assurance/artifact rejection software *art* (http://www.nitrc.org/projects/artifact_detect); structure-function coregistration using each volunteer's high-resolution T1-weighted image; spatial normalisation to Montreal Neurological Institute (MNI-152) standard space with 2mm isotropic resampling resolution, using the segmented grey matter image , together with an *a priori* grey matter template.

To reduce noise due to cardiac, breathing, and motion artifacts, which are known to impact functional connectivity and network analyses [180,181], we applied the anatomical CompCor method of denoising the functional data [182], also implemented within the CONN toolbox. As for preprocessing, we followed the same denoising described in previous work [53,122,175]. The anatomical CompCor method involves regressing out of the functional data the following confounding effects: the first five principal components attributable to each individual's white matter signal, and the first five components attributable to individual cerebrospinal fluid (CSF) signal; six subject-specific realignment parameters (three translations and three rotations) as well as their first- order temporal derivatives; the artefacts identified by *art;* and main effect of scanning condition [182]. Linear detrending was also applied, and the subject-specific denoised BOLD signal timeseries were band-pass filtered to eliminate both low-frequency drift effects and high-frequency noise, thus retaining temporal frequencies between 0.008 and 0.09 Hz.

The step of global signal regression (GSR) has received substantial attention in the fMRI literature, as a potential denoising step [183–186]. However, GSR mathematically mandates that approximately 50% of correlations between regions will be negative [186], thereby removing potentially meaningful differences in the proportion of anticorrelations; additionally, it has been shown across species and states of consciousness that the global signal contains information relevant for consciousness [187]. Therefore, here we chose to avoid GSR in favour of the aCompCor denoising procedure, in line with previous work [53,122,175].

Due to the presence of deformations caused by brain injury, rather than relying on automated pipelines, DOC patients' brains were individually preprocessed using SPM12, with visual inspections after each step. Additionally, to further reduce potential movement artefacts, data underwent despiking with a hyperbolic tangent squashing function, also implemented from the CONN toolbox [179]. The remaining preprocessing and denoising steps were the same as described above.

Brain Parcellation

Brains were parcellated into 454 cortical and subcortical regions of interest (ROIs). The 400 cortical ROIs were obtained from the scale-400 version of the recent Schaefer local-global functional parcellation [62]. Since this parcellation only includes cortical regions, it was augmented with 54 subcortical ROIs from the highest resolution of the recent Tian parcellation [64]. We refer to this 454-ROI parcellation as the "augmented Schaefer" [61]. To ensure the robustness of our results to the choice of atlas, we also replicated them using an alternative cortical parcellation of different dimensionality: we used the Schaefer scale-200 cortical parcellation, complemented with the scale-32 subcortical ROIs from the Tian subcortical atlas [61]. The timecourses of denoised BOLD signals were averaged between all voxels belonging to a given atlas-derived ROI, using the CONN toolbox. The resulting region-specific timecourses of each subject were then extracted for further analysis in MATLAB.

HRF deconvolution

In accordance with our previous work [50,52] and previous studies using of information-theoretic measures in the context of functional MRI data, we used a dedicated toolbox [110] to deconvolve the hemodynamic response function from our regional BOLD signal timeseries prior to analysis.

Measuring Integrated Information

The framework of integrated information decomposition (ΦID) unifies integrated information theory (IIT) and partial information decomposition (PID) to decompose information flow into interpretable, disjoint parts. In this section we provide a brief description of ΦID and formulae required to compute the results. For further details, see [49,50].

Partial information decomposition

We begin with Shannon's Mutual information (MI), which quantifies the interdependence between two random variables *X* and *Y*. It is calculated as

$$\mathrm{I(X;Y)} = \mathrm{H(X)} - \mathrm{H(X|Y)} = \mathrm{H(X)} + \mathrm{H(Y)} - \mathrm{H(X,Y)} \tag{1}$$

where H(*X*) stands for the Shannon entropy of a variable *X*. Above, the first equality states that the mutual information is equal to the reduction in entropy (i.e., uncertainty) about *X* after *Y* is known. Put simply, the mutual information quantifies the information that one variable provides about another [188].

Crucially, Williams and Beer [33] observed that the information that two source variables *X* and *Y* give about a third target variable *Z*, I(*X,Y* ; *Z*), should be decomposable in terms of different *types* of information: information provided by one source but not the other (unique information), by both sources separately (redundant information), or jointly by their combination (synergistic information). Following this intuition, they developed the *Partial Information Decomposition* (PID; [33]) framework, which leads to the following fundamental decomposition:

$$\mathrm{I(X,Y;Z)} = \mathrm{Red(X,Y;Z)} + \mathrm{Un(X;Z\backslash Y)} + \mathrm{Un(Y;Z\backslash X)} + \mathrm{Syn(X,Y;Z)}. \tag{2}$$

Above, *Un* corresponds to the unique information one source but the other doesn't, *Red* is the redundancy between both sources, and *Syn* is their synergy: information that neither *X* nor *Y* alone can provide, but that can be obtained by considering *X* and *Y* together.

The simplest example of a purely synergistic system is one in which *X* and *Y* are independent fair coins, and *Z* is determined by the exclusive-OR function *Z* = XOR(*X,Y*): i.e., *Z*=0 whenever *X* and *Y* have the same value, and *Z*=1 otherwise. It can be shown that *X* and *Y* are both statistically independent of *Z*, which implies that neither of them provide - by themselves - information about *Z*. However, *X* and *Y* together fully determine *Z*, hence the relationship between *Z* with *X* and *Y* is purely synergistic.

As another example for the case of Gaussian variables (as employed here), consider a 2-node coupled autoregressive process with two parameters: a noise correlation *c* and a coupling parameter *a*. As *c* increases, the system is flooded by "common noise", making the system increasingly redundant because the common noise "swamps" the signal of each node. As *a* increases, each node has a stronger influence both on the other and on the system as a whole, and we expect synergy to increase. Therefore, synergy reflects the joint contribution of parts of the system to the whole that is not driven by common noise. This has been demonstrated through computational modelling [48].

Recently, Mediano *et al* (2021) [49] formulated an extension of PID able to decompose the information that multiple source variables have about multiple target variables. This makes PID applicable to the dynamical systems setting, and yields a decomposition with redundant, unique, and synergistic components in the past and future that can be used as a principled method to analyse information flow in neural activity (Figure 3).

Synergy and redundancy calculation

While there is ongoing research on the advantages of different information decompositions for discrete data, most decompositions converge into the same simple form for the case of continuous Gaussian variables [189]. Known as *minimum mutual information PID* (MMI-PID), this decomposition quantifies redundancy in terms of the minimum mutual information of each individual source with the target; synergy, then, becomes identified with the additional information provided by the weaker source once the stronger source is known. Since linear-Gaussian models are sufficiently good descriptors of functional MRI timeseries (and more complex, non-linear models offer no advantage [190,191]), here we adopt the MMI-PID decomposition, following our own and others' previous applications of PID to neuroscientific data [50].

In a dynamical system such as the brain, one can calculate the amount of information flowing from the system's past to its future, known as time-delayed mutual information (TDMI). Specifically, by denoting the past of variables as $X_{t-\tau}$ and $Y_{t-\tau}$ and treating them as sources, and their joint future state (X_t, Y_t), as target, one can apply the PID framework and decompose the information flowing from past to future as

$$\begin{aligned} \mathrm{I}(\mathrm{X}_{t-\tau}, \mathrm{Y}_{t-\tau};\ \mathrm{X}_t, \mathrm{Y}_t) \\ = \mathrm{Red}(\mathrm{X}_{t-\tau}, \mathrm{Y}_{t-\tau};\ \mathrm{X}_t, \mathrm{Y}_t) + \mathrm{Un}(\mathrm{X}_{t-\tau}; \mathrm{X}_t, \mathrm{Y}_t | \mathrm{Y}_{t-\tau}) + \mathrm{Un}(\mathrm{Y}_{t-\tau}; \mathrm{X}_t, \mathrm{Y}_t \backslash \mathrm{X}_{t-\tau}) \\ + \mathrm{Syn}(\mathrm{X}_{t-\tau}, \mathrm{Y}_{t-\tau};\ \mathrm{X}_t, \mathrm{Y}_t) \end{aligned} \tag{3}$$

Applying ΦID to this quantity allows us to distinguish between redundant, unique, and synergistic information shared with respect to the future variables X_t, Y_t [49,50]. Importantly, this framework, has identified $Syn(X_{t-\tau}, Y_{t-\tau};\ X_t, Y_t)$ with the capacity of the system to exhibit emergent behaviour [192] as well as a stronger notion of redundancy, in which information is shared by *X* and *Y* in both past and future. Accordingly, using the MMI-ΦID decomposition for Gaussian variables, we use

$$\mathrm{Red}(\mathrm{X}, \mathrm{Y}) = \min\{\mathrm{I}(\mathrm{X}_{t-\tau};\ \mathrm{X}_t), \mathrm{I}(\mathrm{X}_{t-\tau};\ \mathrm{Y}_t), \mathrm{I}(\mathrm{Y}_{t-\tau};\ \mathrm{X}_t), \mathrm{I}(\mathrm{Y}_{t-\tau};\ \mathrm{Y}_t)\} \tag{4}$$

$$\mathrm{Syn}(\mathrm{X}, \mathrm{Y}) = \mathrm{I}(\mathrm{X}_{t-\tau}, \mathrm{Y}_{t-\tau};\ \mathrm{X}_t, \mathrm{Y}_t) - \max\{\mathrm{I}(\mathrm{X}_{t-\tau};\ \mathrm{X}_t, \mathrm{Y}_t), \mathrm{I}(\mathrm{Y}_{t-\tau};\ \mathrm{X}_t, \mathrm{Y}_t)\}. \tag{5}$$

Here, we used the Gaussian solver implemented in the JIDT toolbox [193] to obtain TDMI, synergy and redundancy between each pair of brain regions, based on their HRF-deconvolved BOLD signal timeseries" [49,50].

Revised measure of integrated information from Integrated Information Decomposition

Through the framework of Integrated Information Decomposition, we can decompose the constituent elements of Φ, the formal measure of integrated information proposed by Integrated Information Theory to quantify consciousness [31]. Note that several variants of Φ have been proposed over the years, including the original formulation of Tononi [29], other formulations based on causal perturbation [113,194] and others (see [48,195] for comparative reviews). Here, we focus on the "empirical Φ" measure of Seth and Barrett [196], based on the

measures by Balduzzi and Tononi (2008) [31] and adapted to applications to experimental data. It is computed as

$$\Phi = \mathrm{I}(\mathrm{X}_{t-\tau}, \mathrm{Y}_{t-\tau};\ \mathrm{X}_t, \mathrm{Y}_t) - \mathrm{I}(\mathrm{X}_{t-\tau};\ \mathrm{X}_t) - \mathrm{I}(\mathrm{Y}_{t-\tau};\ \mathrm{Y}_t) \tag{6}$$

and it quantifies how much temporal information is contained in the system over and above the information in its past. This measure is easy to compute (compared with other Φ measures) [197] and represents a noteworthy attempt to formalise the powerful intuitions underlying IIT. However, once the original formulation from Balduzzi and Tononi is rendered suitable for practical empirical application [196,198] the resulting mathematical formulation has known shortcomings, including the fact that it can yield negative values in some cases - which are hard to interpret and seemingly paradoxical, as it does not seem plausible for a system to be "negatively integrated" or an organism to have negative consciousness [196,198].

Interestingly, with ΦID it can be formally demonstrated [49] that Φ is composed of different information atoms: it contains all the synergistic information in the system, the unique information transferred from X to Y and vice versa, and, importantly, the subtraction of redundancy - which explains why Φ can be negative in redundancy-dominated systems.

To address this fundamental shortcoming, Mediano, Rosas and colleagues [49] introduced a revised measure of integrated information, Φ_R, which consists of the original Φ with the redundancy added back in:

$$\Phi_R = \Phi + \mathrm{Red}(\mathrm{X}, \mathrm{Y}) \tag{7}$$

where Red(*X, Y*) is defined in Eq. (4). This measure is computationally tractable and preserves the original intuition of integrated information as measuring the extent to which "the whole is greater than the sum of its parts", since it captures only synergistic and transferred information. Crucially, thanks to Integrated Information Decomposition, it can be proved that the improved formulation of integrated information that we adopt here is guaranteed to be non-negative [49] - thereby avoiding a major conceptual limitation of the original formulation of Φ.

Note that the formula for Φ^{WMS} above stems from what is known as IIT 2.0, but TDMI is by no means the only way of quantifying the dynamical structure of a system: indeed, subsequent developments in IIT 3.0 used alternative metrics with a more explicit focus on causal interpretations [113], which were in turn replaced in the latest iteration known as IIT 4.0 [194,199]. We do not consider the alternative measure of integrated information proposed in IIT 3.0 because it is computationally intractable for systems bigger than a small set of logic gates, and it is not universally well-defined [198].

Gradient of redundancy-to-synergy relative importance to identify the synergistic workspace

After building networks of synergistic and redundant interactions between each pair of regions of interest (ROIs), we determined the role of each ROI in terms of its relative engagement in synergistic or redundant interactions. Following the procedure previously described [50], we first calculated the nodal strength of each brain region as the sum of all its interactions in the group-averaged matrix (Figure S1). Then, we ranked all 454 regions based on their nodal strength (with higher-strength regions having higher ranks). This procedure was done separately for networks of synergy and redundancy. Subtracting each region's redundancy rank from its synergy rank yielded a gradient from negative (i.e.,

ranking higher in terms of redundancy than synergy) to positive (i.e., having a synergy rank higher than the corresponding redundancy rank; note that the sign is arbitrary).

It is important to note that the gradient is based on relative - rather than absolute - differences between regional synergy and redundancy; consequently, a positive rank difference does not necessarily mean that the region's synergy is greater than its redundancy; rather, it indicates that the balance between its synergy and redundancy relative to the rest of the brain is in favour of synergy - and *vice versa* for a negative gradient.

Subdivision of workspace nodes into gateways and broadcasters

To identify which regions within the workspace play the role of gateways or broadcasters postulated in our proposed architecture, we followed a procedure analogous to the one adopted to identify the gradient of redundancy-synergy relative importance, but replacing the node *strength* with the node *participation coefficient*. The participation coefficient P_i quantifies the degree of connection that a node entertains with nodes belonging to other modules: the more of a node's connections are towards other modules, the higher its participation coefficient will be [60,200]. Conversely, the participation coefficient of a node will be zero if its connections are all with nodes belonging to its own module.

$$P_i = 1 - \sum_{s=1}^{M} \left(\frac{\kappa_{is}}{k_i}\right)^2 \tag{8}$$

Here, κ_{is} is the strength of positive connections between node *i* and other nodes in module *s*, k_i is the strength of all its positive connections, and *M* is the number of modules in the network. The participation coefficient ranges between zero (no connections with other modules) and one (equal connections to all other modules) [60,200].

Here, modules were set to be the seven canonical resting-state networks identified by Yeo and colleagues [63], into which the Schaefer parcellation is already divided [62], with the addition of an eighth subcortical network comprising all ROIs of the Tian subcortical network [64]. The brain's RSNs were chosen as modules because of their distinct and well-established functional roles, which fit well with the notion of modules as segregated and specialised processing systems interfacing with the global workspace. Additionally, having the same definition of modules (i.e., RSNs) for synergy and redundancy allowed us to compute their respective participation coefficients in an unbiased way.

Separately for connectivity matrices of synergy and redundancy, the participation coefficient of each brain region was calculated. Then, regions belonging to the synergistic workspace were ranked, so that higher ranks indicated higher participation coefficient. Finally, the redundancy-based participation coefficient rank of each workspace region was subtracted from its corresponding synergy-based participation coefficient rank, to quantify – within the workspace – whether regions have relatively more diverse connectivity in terms of synergy, or in terms of redundancy.

This procedure yielded a gradient over workspace regions, from negative (i.e. having a more highly ranked participation coefficient based on redundancy than synergy) to positive (i.e. having a more highly ranked participation coefficient based on synergy than redundancy). Note that as before, the sign of this gradient is arbitrary, and it is based on relative rather than absolute difference. Workspace regions with a positive gradient value were classified as "gateways", since they have synergistic interactions with many brain modules. In contrast, workspace regions with a negative value of the gradient - i.e. those whose redundancy rank is higher than their synergy rank, in terms of participation coefficient - were labelled as

workspace "broadcasters", since they possess information that is duplicated across multiple modules in the brain.

Statistical Analysis

Network Based Statistic

The network-based statistic approach [67] was used to investigate the statistical significance of propofol-induced or DOC-induced alterations. This nonparametric statistical method is designed to control the family-wise error due to multiple comparisons, for application to graph data. Connected components of the graph are identified from edges that survive an a-priori statistical threshold (F-contrast; here we set the threshold to an F-value of 9, two-sided, with an alpha level of 0.05). In turn, the statistical significance of such connected components is estimated by comparing their topology against a null distribution of the size of connected components obtained from non-parametric permutation testing. This approach rejects the null hypothesis on a component-by-component level, and therefore achieves superior power compared to mass-univariate approaches [67].

Testing for shared effects across datasets

We sought to detect changes that are common across datasets, to rule out possible propofol- or DOC-specific effects that are not related to consciousness per se [53]. To this end, we employed a null hypothesis significance test under the composite null hypothesis that *at least* one dataset among those considered here has no effect. In other words, for the null hypothesis to be rejected we demand that all comparisons exhibit non-zero effects. As usual, the test proceeds by comparing an observed test statistic with a null distribution. The test statistic is the minimum of the three F-scores obtained in the comparisons of interest (DOC vs awake; anaesthesia vs awake; and anaesthesia vs recovery), and the null distribution is sampled by randomly reshuffling exactly one dataset (picked at random) at a time and recalculating the F-scores. By shuffling exactly one dataset (instead of all of them), we are comparing the observed data against the "least altered" version of the data that is still compatible with the null hypothesis. This is a type of least favourable configuration test [201], which is guaranteed to control the false positive rate below a set threshold (here, 0.05). The details of this test will be described in a future publication. Common changes across the three states of consciousness were then identified as edges (defined in terms of Φ_R) that were either (i) increased in DOC compared with control; (ii) increased during anaesthesia compared with wakefulness; and (iii) increased during anaesthesia compared with post-anaesthetic recovery; or (i) decreased in DOC compared with control; (ii) decreased during anaesthesia compared with wakefulness; and (iii) decreased during anaesthesia compared with post-anaesthetic recovery.

Spatial autocorrelation-preserving null model for correlation

The significance of correlation between nodes' participation coefficient based on different definitions of modules (a-priori as resting-state networks or in a data-driven fashion from Louvain community detection) was assessed using a spatial permutation test which generates a null distribution of 10,000 randomly rotated brain maps with preserved spatial covariance ('spin test'), to ensure robustness to the potential confounding effects of spatial autocorrelation [202–204].

Data Availability

The raw data analysed during the current study are available on request from the following authors. Propofol anaesthesia, Disorders of Consciousness and test-retest datasets: Dr. Emmanuel A. Stamatakis (University of Cambridge, Division of Anaesthesia; email:

eas46@cam.ac.uk). The Human Connectome Project datasets are freely available from http://www.humanconnectome.org/.

Code availability

The Java Information Dynamics Toolbox v1.5 is freely available online: (https://github.com/jlizier/jidt). The CONN toolbox version 17f is freely available online (http://www.nitrc.org/projects/conn). The Brain Connectivity Toolbox code used for graph-theoretical analyses is freely available online (https://sites.google.com/site/bctnet/). The HRF deconvolution toolbox v2.2 is freely available online: (https://www.nitrc.org/projects/rshrf). The code for spin-based permutation testing of cortical correlations is freely available at https://github.com/frantisekvasa/rotate_parcellation. We have made freely available MATLAB/Octave and Python code to compute measures of Integrated Information Decomposition of timeseries with the Gaussian MMI solver, at https://github.com/Imperial-MIND-lab/integrated-info-decomp.

Acknowledgments

Author AIL is grateful to Dr. Athena Demertzi and Dr. Petra Vertes for helpful discussion. This work was supported by grants from the UK Medical Research Council [U.1055.01.002.00001.01 to AMO and JDP]; The James S. McDonnell Foundation [to AMO and JDP]; and the Canada Excellence Research Chairs program (215063 to AMO); the National Institute for Health Research (NIHR, UK), Cambridge Biomedical Research Centre and NIHR Senior Investigator Awards [to DKM], the Stephen Erskine Fellowship (Queens' College, Cambridge, to EAS), the L'Oreal-Unesco for Women in Science Excellence Research Fellowship to LN; the British Oxygen Professorship of the Royal College of Anaesthetists [to DKM] and the Gates Cambridge Trust (OPP 1144; to AIL). PAM and DB are funded by the Wellcome Trust (grant no. 210920/Z/18/Z). FR is funded by the Ad Astra Chandaria foundation. The research was also supported by the NIHR Brain Injury Healthcare Technology Co-operative based at Cambridge University Hospitals NHS Foundation Trust and University of Cambridge. AMO and DKM are Fellows of the CIFAR Brain, Mind, and Consciousness Programme. Data were provided [in part] by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

Author Contributions: AIL: conceived the study; analysed data; wrote first draft of the manuscript. PAM: conceived the study; contributed to data analysis and interpretation of results; reviewed and edited the manuscript. FR: contributed to data analysis and interpretation of results; reviewed and edited the manuscript. M.M.C.: contributed to data analysis. DKM: reviewed the manuscript. RCH: reviewed the manuscript and contributed to interpretation of results. DB: conceived the study; reviewed and edited the manuscript. EAS: conceived the study; reviewed and edited the manuscript. P.F., G.B.W., J.A., J.D.P., A.M.O., L.N., D.K.M. and E.A.S. were involved in designing the original studies for which the present data were collected. P.F., M.M.C., G.B.W., J.A., L.N. and E.A.S. all participated in data collection.

Competing Interest Statement: The authors declare no competing interests.

References

1. Varela, F., Lachaux, J.P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: Phase synchronization and large-scale integration. Nature Reviews Neuroscience *2*, 229–239. 10.1038/35067550.

2. Petersen, S.E., and Sporns, O. (2015). Brain Networks and Cognitive Architectures. Neuron *88*, 207–219. 10.1016/j.neuron.2015.09.027.

3. Shine, J.M. (2019). Neuromodulatory Influences on Integration and Segregation in the Brain. Trends in Cognitive Sciences *23*, 572–583. 10.1016/j.tics.2019.04.002.

4. Fodor, J.A. (1985). Précis of *The Modularity of Mind*. Behavioral and Brain Sciences *8*, 1–5. 10.1017/S0140525X0001921X.

5. Deco, G., Vidaurre, D., and Kringelbach, M.L. (2021). Revisiting the global workspace: Orchestration of the functional hierarchical organisation of the human brain. Nature Human Behaviour *5*. 10.1101/859579.

6. Assem, M., Glasser, M.F., Van Essen, D.C., and Duncan, J. (2020). A Domain-General Cognitive Core Defined in Multimodally Parcellated Human Cortex. Cerebral Cortex *30*, 4361–4380. 10.1093/cercor/bhaa023.

7. Shine, J.M., Breakspear, M., Bell, P.T., Ehgoetz Martens, K., Shine, R., Koyejo, O., Sporns, O., and Poldrack, R.A. (2019). Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. Nature Neuroscience *22*, 289–296. 10.1038/s41593-018-0312-0.

8. Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., and Honey, C.J. (2008). Mapping the structural core of human cerebral cortex. PLoS Biol *6*, 159. 10.1371/journal.pbio.0060159.

9. Felleman, D.J., and Van Essen, D.C. (1991). Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex *1*, 1–47. 10.1093/cercor/1.1.1.

10. Goulas, A., Changeux, J.P., Wagstyl, K., Amunts, K., Palomero-Gallagher, N., and Hilgetag, C.C. (2021). The natural axis of transmitter receptor distribution in the human cerebral cortex. Proceedings of the National Academy of Sciences of the United States of America *118*. 10.1073/PNAS.2020574118.

11. Sydnor, V.J., Larsen, B., Bassett, D.S., Alexander-Bloch, A., Fair, D.A., Liston, C., Mackey, A.P., Milham, M.P., Pines, A., Roalf, D.R., et al. (2021). Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. Neuron *109*, 2820–2846. 10.1016/j.neuron.2021.06.016.

12. Baum, G.L., Cui, Z., Roalf, D.R., Ciric, R., Betzel, R.F., Larsen, B., Cieslak, M., Cook, P.A., Xia, C.H., Moore, T.M., et al. (2020). Development of structure–function coupling in human brain networks during youth. Proceedings of the National Academy of Sciences of the United States of America *117*, 771–778. 10.1073/pnas.1912034117.

13. Bertolero, M.A., Yeo, B.T.T., and D'Esposito, M. (2017). The diverse club. Nature Communications *8*, 1277. 10.1038/s41467-017-01189-w.

1195 14. Vázquez-Rodríguez, B., Suárez, L.E., Markello, R.D., Shafiei, G., Paquola, C., Hagmann, P., Van
1196 Den Heuvel, M.P., Bernhardt, B.C., Spreng, R.N., and Misic, B. (2019). Gradients of structure–
1197 function tethering across neocortex. Proceedings of the National Academy of Sciences of the
1198 United States of America *116*, 21219–21227. 10.1073/pnas.1903403116.

1199 15. Margulies, D.S., Ghosh, S.S., Goulas, A., Falkiewicz, M., Huntenburg, J.M., Langs, G., Bezgin, G.,
1200 Eickhoff, S.B., Castellanos, F.X., Petrides, M., et al. (2016). Situating the default-mode network
1201 along a principal gradient of macroscale cortical organization. Proceedings of the National
1202 Academy of Sciences of the United States of America *113*, 12574–12579.
1203 10.1073/pnas.1608282113.

1204 16. Burt, J.B., Demirtaş, M., Eckner, W.J., Navejar, N.M., Ji, J.L., Martin, W.J., Bernacchia, A.,
1205 Anticevic, A., and Murray, J.D. (2018). Hierarchy of transcriptomic specialization across human
1206 cortex captured by structural neuroimaging topography. Nature Neuroscience *21*, 1251–1259.
1207 10.1038/s41593-018-0195-0.

1208 17. Demirtaş, M., Burt, J.B., Helmer, M., Ji, J.L., Adkinson, B.D., Glasser, M.F., Van Essen, D.C.,
1209 Sotiropoulos, S.N., Anticevic, A., and Murray, J.D. (2019). Hierarchical Heterogeneity across
1210 Human Cortex Shapes Large-Scale Neural Dynamics. Neuron *101*, 1181-1194.e13.
1211 10.1016/j.neuron.2019.01.017.

1212 18. Deco, G., Kringelbach, M.L., Arnatkeviciute, A., Oldham, S., Sabaroedin, K., Rogasch, N.C.,
1213 Aquino, K.M., and Fornito, A. (2021). Dynamical consequences of regional heterogeneity in the
1214 brain's transcriptional landscape. Science Advances *7*, eabf4752. 10.1126/SCIADV.ABF4752.

1215 19. Hansen, J.Y., Markello, R.D., Vogel, J.W., Seidlitz, J., Bzdok, D., and Misic, B. (2021). Mapping
1216 gene transcription and neurocognition across human neocortex. Nature Human Behaviour *5*,
1217 1240–1250. 10.1038/s41562-021-01082-z.

1218 20. Hansen, J.Y., Shafiei, G., Markello, R.D., Smart, K., Cox, S., Norgaard, M., Beliveau, V., Wu, Y.,
1219 Gallezot, J.-D., Aumont, E., et al. (2022). Mapping neurotransmitter systems to the structural
1220 and functional organization of the human neocortex. Nature Neuroscience *25*, 1569–1581.

1221 21. Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: From
1222 consciousness to its physical substrate. Nature Reviews Neuroscience *17*, 450–461.
1223 10.1038/nrn.2016.44.

1224 22. Seth, A.K., and Bayne, T. (2022). Theories of consciousness. Nature reviews. Neuroscience.
1225 10.1038/S41583-022-00587-4.

1226 23. Dehaene, S., and Changeux, J.-P. (2011). Experimental and Theoretical Approaches to
1227 Conscious Processing. Neuron *70*, 200–227. 10.1016/j.neuron.2011.03.018.

1228 24. Mashour, G.A., Roelfsema, P., Changeux, J.P., and Dehaene, S. (2020). Conscious Processing
1229 and the Global Neuronal Workspace Hypothesis. Neuron *105*, 776–798.
1230 10.1016/j.neuron.2020.01.026.

1231 25. Dehaene, S., Changeux, J.P., and Naccache, L. (2011). The global neuronal workspace model of
1232 conscious access: From neuronal architectures to clinical applications. Research and
1233 Perspectives in Neurosciences *18*, 55–84. 10.1007/978-3-642-18015-6_4.

1234 26. Baars, B.J. (2005). Global workspace theory of consciousness: Toward a cognitive neuroscience
1235 of human experience. Progress in Brain Research *150*, 45–53. 10.1016/S0079-6123(05)50004-
1236 9.

1237 27. Dehaene, S., and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness:
1238 Basic evidence and a workspace framework. Cognition *79*, 1–37. 10.1016/S0010-
1239 0277(00)00123-2.

1240 28. Tononi, G. (2008). Consciousness as Integrated Information. The Biological Bulletin *215*, 216–
1241 242. 215/3/216 [pii].

1242 29. Tononi, G. (2004). An information integration theory of consciousness. BMC Neuroscience *5*,
1243 42–64. 10.1186/1471-2202-5-42.

1244 30. Tononi, G., Edelman, G.M., and Sporns, O. (1998). Complexity and coherency: Integrating
1245 information in the brain. Trends in Cognitive Sciences *2*, 474–484. 10.1016/S1364-
1246 6613(98)01259-5.

1247 31. Balduzzi, D., and Tononi, G. (2008). Integrated information in discrete dynamical systems:
1248 Motivation and theoretical framework. PLoS Computational Biology *4*.
1249 10.1371/journal.pcbi.1000091.

1250 32. Cavanna, F., Vilas, M.G., Palmucci, M., and Tagliazucchi, E. (2018). Dynamic functional
1251 connectivity and brain metastability during altered states of consciousness. NeuroImage *180*,
1252 383–395. 10.1016/j.neuroimage.2017.09.065.

1253 33. Williams, P.L., and Beer, R.D. (2010). Nonnegative Decomposition of Multivariate Information.
1254 arXiv, http://arxiv.org/abs/1004.2515.

1255 34. Luppi, A.I., Rosas, F.E., Mediano, P.A.M., Menon, D.K., and Stamatakis, E.A. (2024). Information
1256 decomposition and the informational architecture of the brain. Trends in Cognitive Sciences *0*.
1257 10.1016/j.tics.2023.11.005.

1258 35. Quian Quiroga, R., and Panzeri, S. (2009). Extracting information from neuronal populations:
1259 Information theory and decoding approaches. Nature Reviews Neuroscience *10*, 173–185.
1260 10.1038/nrn2578.

1261 36. Timme, N., Alford, W., Flecker, B., and Beggs, J.M. (2014). Synergy, redundancy, and
1262 multivariate information measures: An experimentalist's perspective. Journal of Computational
1263 Neuroscience *36*, 119–140. 10.1007/s10827-013-0458-4.

1264 37. Sherrill, S.P., Timme, N., Beggs, J., and Newman, E.L. (2021). Partial information decomposition
1265 reveals that synergistic neural integration is greater downstream of recurrent information flow
1266 in organotypic cortical cultures. 1–28. 10.1371/journal.pcbi.1009196.

1267 38. Faber, S.P., Timme, N.M., Beggs, J.M., and Newman, E.L. (2019). Computation is concentrated
1268 in rich clubs of local cortical networks. Network Neuroscience *3*, 384–404.
1269 10.1162/netn_a_00069.

1270 39. Sherrill, S.P., Timme, N.M., Beggs, J.M., and Newman, E.L. (2020). Correlated activity favors
1271 synergistic processing in local cortical networks in vitro at synaptically relevant timescales.
1272 Network Neuroscience *4*, 678–697. 10.1162/netn_a_00141.

40. Newman, E.L., Varley, T.F., Parakkattu, V.K., Sherrill, S.P., and Beggs, J.M. (2022). Revealing the Dynamics of Neural Information Processing with Multivariate Information Decomposition. Entropy *24*, 930. 10.3390/e24070930.

41. Celotto, M., Bim, J., Tlaie, A., de Feo, V., Toso, A., Lemke, S., Chicharro, D., Nili, H., Bieler, M., Hanganu-Opatz, I.L., et al. (2023). An information-theoretic quantification of the content of communication between brain regions Poster. In NeurIPS.

42. Varley, T.F. (2023). Decomposing past and future: Integrated information decomposition based on shared probability mass exclusions. PLOS ONE *18*, e0282950. 10.1371/journal.pone.0282950.

43. Varley, T.F., Sporns, O., Schaffelhofer, S., Scherberger, H., and Dann, B. (2023). Information-processing dynamics in neural networks of macaque cerebral cortex reflect cognitive state and behavior. Proc Natl Acad Sci U S A *120*, e2207677120. 10.1073/pnas.2207677120.

44. Delis, I., Ince, R.A.A., Sajda, P., and Wang, Q. (2022). Neural Encoding of Active Multi-Sensing Enhances Perceptual Decision-Making via a Synergistic Cross-Modal Interaction. J. Neurosci. *42*, 2344–2355. 10.1523/JNEUROSCI.0861-21.2022.

45. Francis, N.A., Mukherjee, S., Koçillari, L., Panzeri, S., Babadi, B., and Kanold, P.O. (2022). Sequential transmission of task-relevant information in cortical neuronal networks. Cell Reports *39*, 110878. 10.1016/j.celrep.2022.110878.

46. Barnett, L., Barrett, A.B., and Seth, A.K. (2009). Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables. Phys. Rev. Lett. *103*, 238701. 10.1103/PhysRevLett.103.238701.

47. Wibral, M., Lizier, J.T., Vögler, S., Priesemann, V., and Galuske, R. (2014). Local active information storage as a tool to understand distributed neural information processing. Frontiers in Neuroinformatics *8*, 1. 10.3389/FNINF.2014.00001/BIBTEX.

48. Mediano, P.A.M., Seth, A.K., and Barrett, A.B. (2019). Measuring integrated information: Comparison of candidate measures in theory and simulation. Entropy *21*. 10.3390/e21010017.

49. Mediano, P.A.M., Rosas, F.E., Luppi, A.I., Carhart-Harris, R.L., Bor, D., Seth, A.K., and Barrett, A.B. (2021). Towards an extended taxonomy of information dynamics via Integrated Information Decomposition. arXiv.

50. Luppi, A.I., Mediano, P.A.M., Rosas, F.E., Holland, N., Fryer, T.D., O'Brien, J.T., Rowe, J.B., Menon, D.K., Bor, D., and Stamatakis, E.A. (2022). A synergistic core for human brain evolution and cognition. Nat Neurosci *25*, 771–782. 10.1038/s41593-022-01070-0.

51. Varley, T.F., Pope, M., Puxeddu, M.G., Faskowitz, J., and Sporns, O. (2023). Partial entropy decomposition reveals higher-order structures in human brain activity. PNAS *120*. https://doi.org/10.1073/pnas.2300888120.

52. Luppi, A.I., Mediano, P.A.M., Rosas, F.E., Allanson, J., Pickard, J.D., Williams, G.B., Craig, M.M., Finoia, P., Peattie, A.R.D., Coppola, P., et al. (2023). Reduced emergent character of neural dynamics in patients with a disrupted connectome. Neuroimage, 119926. 10.1016/j.neuroimage.2023.119926.

53. Luppi, A.I., Craig, M.M., Pappas, I., Finoia, P., Williams, G.B., Allanson, J., Pickard, J.D., Owen, A.M., Naci, L., Menon, D.K., et al. (2019). Consciousness-specific dynamic interactions of brain integration and functional diversity. Nature Communications *10*. 10.1038/s41467-019-12658-9.

54. Tsitsiklis, J.N. (1989). Decentralized Detection. In Advances in Statistical Signal Processing, Vol 2: Signal Detection.

55. Veeravalli, V.V., and Varshney, P.K. (2012). Distributed inference in wireless sensor networks. Trans. R. Soc. A *370*, 100–117. 10.1098/rsta.2011.0194.

56. Lever, K.E., Merabti, M., and Kifayat, K. (2013). Single Points of Failure Within Systems-of-Systems. 14th Annual Post Graduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting (PGNet), 183–188.

57. Rosas, F., Hsiao, J.H., and Chen, K.C. (2017). A technological perspective on information cascades via social learning. IEEE Access *5*, 22605–22633. 10.1109/ACCESS.2017.2687422.

58. Sneve, M.H., Grydeland, H., Rosa, M.G.P., Paus, T., Chaplin, T., Walhovd, K., and Fjell, A.M. (2019). High-expanding regions in primate cortical brain evolution support supramodal cognitive flexibility. Cerebral Cortex *29*, 3891–3901. 10.1093/cercor/bhy268.

59. Shanahan, M. (2012). The brain's connective core and its role in animal cognition. Philosophical Transactions of the Royal Society B: Biological Sciences *367*, 2704–2714. 10.1098/rstb.2012.0128.

60. Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. NeuroImage *52*, 1059–1069. 10.1016/j.neuroimage.2009.10.003.

61. Luppi, A.I., and Stamatakis, E.A. (2021). Combining network topology and information theory to construct representative brain networks. Network Neuroscience *5*, 96–124. 10.1162/netn_a_00170.

62. Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.-N., Holmes, A.J., Eickhoff, S.B., and Yeo, B.T.T. (2018). Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. Cerebral Cortex *28*, 3095–3114. 10.1093/cercor/bhx179.

63. Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zollei, L., Polimeni, J.R., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of neurophysiology *106*, 1125–1165. 10.1152/jn.00338.2011.

64. Tian, Y., Margulies, D.S., Breakspear, M., and Zalesky, A. (2020). Topographic organization of the human subcortex unveiled with functional connectivity gradients. Nature Neuroscience *23*, 1421–1432. 10.1038/s41593-020-00711-6.

65. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment *2008*, P10008. 10.1088/1742-5468/2008/10/P10008.

66. Bor, D., and Seth, A.K. (2012). Consciousness and the prefrontal parietal network: Insights from attention, working memory, and chunking. Frontiers in Psychology *3*. 10.3389/fpsyg.2012.00063.

1351 67. Zalesky, A., Fornito, A., and Bullmore, E.T. (2010). Network-based statistic: Identifying
1352 differences in brain networks. NeuroImage *53*, 1197–1207.
1353 10.1016/j.neuroimage.2010.06.041.

1354 68. Fedorenko, E., Duncan, J., and Kanwisher, N. (2013). Broad domain generality in focal regions
1355 of frontal and parietal cortex. Proceedings of the National Academy of Sciences of the United
1356 States of America *110*, 16616–16621. 10.1073/pnas.1315235110.

1357 69. Duncan, J., and Owen, A.M. (2000). Common regions of the human frontal lobe recruited by
1358 diverse cognitive demands. Trends in Neurosciences *23*, 475–483. 10.1016/S0166-
1359 2236(00)01633-7.

1360 70. Barbey, A.K. (2018). Network Neuroscience Theory of Human Intelligence. Trends in Cognitive
1361 Sciences *22*, 8–20. 10.1016/j.tics.2017.10.001.

1362 71. Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q.K., Yu, A.B., Kahn, A.E., Medaglia, J.D., Vettel,
1363 J.M., Miller, M.B., Grafton, S.T., et al. (2015). Controllability of structural brain networks.
1364 Nature Communications *6*, 8414. 10.1038/ncomms9414.

1365 72. Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., and Raichle, M.E. (2005).
1366 The human brain is intrinsically organized into dynamic, anticorrelated functional networks.
1367 Proceedings of the National Academy of Sciences *102*, 9673–9678. 10.1073/pnas.0504136102.

1368 73. Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., and Shulman, G.L.
1369 (2001). A default mode of brain function. Proceedings of the National Academy of Sciences of
1370 the United States of America *98*, 676–682. 10.1073/pnas.98.2.676.

1371 74. Raichle, M.E. (2015). The Brain's Default Mode Network. Annual Review of Neuroscience *38*,
1372 433–447. 10.1146/annurev-neuro-071013-014030.

1373 75. Wei, Y., de Lange, S.C., Scholtens, L.H., Watanabe, K., Ardesch, D.J., Jansen, P.R., Savage, J.E., Li,
1374 L., Preuss, T.M., Rilling, J.K., et al. (2019). Genetic mapping and evolutionary analysis of human-
1375 expanded cognitive networks. Nature Communications *10*. 10.1038/s41467-019-12764-8.

1376 76. Xu, T., Nenning, K.H., Schwartz, E., Hong, S.J., Vogelstein, J.T., Goulas, A., Fair, D.A., Schroeder,
1377 C.E., Margulies, D.S., Smallwood, J., et al. (2020). Cross-species functional alignment reveals
1378 evolutionary hierarchy within the connectome. NeuroImage *223*.
1379 10.1016/j.neuroimage.2020.117346.

1380 77. Huntenburg, J.M., Bazin, P.-L., and Margulies, D.S. (2018). Large-Scale Gradients in Human
1381 Cortical Organization. Trends Cogn Sci *22*, 21–31. 10.1016/j.tics.2017.11.002.

1382 78. Smallwood, J., Bernhardt, B.C., Leech, R., Bzdok, D., Jefferies, E., and Margulies, D.S. (2021).
1383 The default mode network in cognition: a topographical perspective. Nature Reviews
1384 Neuroscience *22*, 503–513. 10.1038/s41583-021-00474-4.

1385 79. Deco, G., Kringelbach, M.L., Jirsa, V.K., and Ritter, P. (2017). The dynamics of resting
1386 fluctuations in the brain: Metastability and its dynamical cortical core. Scientific Reports *7*, 1–
1387 14. 10.1038/s41598-017-03073-5.

1388 80. Kabbara, A., EL Falou, W., Khalil, M., Wendling, F., and Hassan, M. (2017). The dynamic
1389 functional core network of the human brain at rest. Scientific Reports *7*, 2936.
1390 10.1038/s41598-017-03420-6.

1391 81. de Pasquale, F., Della Penna, S., Snyder, A.Z., Marzetti, L., Pizzella, V., Romani, G.L., and
1392 Corbetta, M. (2012). A Cortical Core for Dynamic Integration of Functional Networks in the
1393 Resting Human Brain. Neuron *74*, 753–764. 10.1016/j.neuron.2012.03.031.

1394 82. Vatansever, D., Menon, X.D.K., Manktelow, A.E., Sahakian, B.J., and Stamatakis, E.A. (2015).
1395 Default Mode Dynamics for Global Functional Integration. The Journal of neuroscience: the
1396 official journal of the Society for Neuroscience *35*, 15254–15262. 10.1523/JNEUROSCI.2135-
1397 15.2015.

1398 83. Vatansever, D., Menon, D.K., and Stamatakis, E.A. (2017). Default mode contributions to
1399 automated information processing. Proceedings of the National Academy of Sciences of the
1400 United States of America *114*, 12821–12826. 10.1073/pnas.1710521114.

1401 84. Chiou, R., Humphreys, G.F., and Lambon Ralph, M.A. (2020). Bipartite functional fractionation
1402 within the default network supports disparate forms of internally oriented cognition. Cerebral
1403 Cortex *30*, 5484–5501. 10.1093/cercor/bhaa130.

1404 85. Cavanna, A.E., and Trimble, M.R. (2006). The precuneus: A review of its functional anatomy
1405 and behavioural correlates. Brain *129*, 564–583. 10.1093/brain/awl004.

1406 86. Qin, P., and Northoff, G. (2011). How is our self related to midline regions and the default-
1407 mode network? NeuroImage *57*, 1221–1233. 10.1016/j.neuroimage.2011.05.028.

1408 87. Karapanagiotidis, T., Bernhardt, B.C., Jefferies, E., and Smallwood, J. (2017). Tracking thoughts:
1409 Exploring the neural architecture of mental time travel during mind-wandering. NeuroImage
1410 *147*, 272–281. 10.1016/j.neuroimage.2016.12.031.

1411 88. Buckner, R.L., Andrews-Hanna, J.R., and Schacter, D.L. (2008). The brain's default network:
1412 Anatomy, function, and relevance to disease. Annals of the New York Academy of Sciences
1413 *1124*, 1–38. 10.1196/annals.1440.011.

1414 89. Buckner, R.L., and DiNicola, L.M. (2019). The brain's default network: updated anatomy,
1415 physiology and evolving insights. Nature Reviews Neuroscience *20*, 593–608. 10.1038/s41583-
1416 019-0212-7.

1417 90. Schacter, D.L., Addis, D.R., and Buckner, R.L. (2007). Remembering the past to imagine the
1418 future: the prospective brain. Nature Reviews Neuroscience *8*, 657–661.
1419 10.1080/08995600802554748.

1420 91. Szpunar, K.K., Spreng, R.N., and Schacter, D.L. (2014). A taxonomy of prospection: Introducing
1421 an organizational framework for future-oriented cognition. Proceedings of the National
1422 Academy of Sciences of the United States of America *111*, 18414–18421.
1423 10.1073/pnas.1417144111.

1424 92. Hassabis, D., and Maguire, E.A. (2009). The construction system of the brain. Philosophical
1425 Transactions of the Royal Society B: Biological Sciences *364*, 1263–1271.
1426 10.1098/rstb.2008.0296.

1427 93. Wen, T., Mitchell, D.J., and Duncan, J. (2020). The Functional Convergence and Heterogeneity
1428 of Social, Episodic, and Self-Referential Thought in the Default Mode Network. Cerebral cortex
1429 *30*, 5915–5929. 10.1093/cercor/bhaa166.

1430 94. Wang, X., Margulies, D.S., Smallwood, J., and Jefferies, E. (2020). A gradient from long-term
1431 memory to novel cognition: Transitions through default mode and executive cortex.
1432 NeuroImage *220*. 10.1016/j.neuroimage.2020.117074.

1433 95. Dohmatob, E., Dumas, G., and Bzdok, D. (2020). Dark control: The default mode network as a
1434 reinforcement learning agent. Human Brain Mapping *41*, 3318–3341. 10.1002/hbm.25019.

1435 96. Yeshurun, Y., Nguyen, M., and Hasson, U. (2021). The default mode network: where the
1436 idiosyncratic self meets the shared social world. Nature Reviews Neuroscience *22*, 181–192.
1437 10.1038/s41583-020-00420-w.

1438 97. Shanahan, M. (2010). Embodiment and the Inner Life: Cognition and Consciousness in the
1439 Space of Possible Minds (Oxford University Press)
1440 10.1093/acprof:oso/9780199226559.001.0001.

1441 98. MacDonald, A.A., Naci, L., MacDonald, P.A., and Owen, A.M. (2015). Anesthesia and
1442 neuroimaging: Investigating the neural correlates of unconsciousness. Trends in Cognitive
1443 Sciences *19*, 100–107. 10.1016/j.tics.2014.12.005.

1444 99. Hannawi, Y., Lindquist, M.A., Caffo, B.S., Sair, H.I., and Stevens, R.D. (2015). Resting brain
1445 activity in disorders of consciousness: a systematic review and meta-analysis. Neurology *84*,
1446 1272–1280. 10.1212/WNL.0000000000001404; 10.1212/WNL.0000000000001404.

1447 100. Vanhaudenhuyse, A., Quentin Noirhomme, Ã., Luaba J-F Tshibanda, Ã., Bruno, M.-A., Boveroux,
1448 P., Schnakers, C., Soddu, A., Perlbarg, V., Ledoux, D., Brichant, J.-F., et al. (2010). Default
1449 network connectivity reflects the level of consciousness in non-communicative brain- damaged
1450 patients. Brain *133*, 161–171. 10.1093/brain/awp313.

1451 101. Di Perri, C., Amico, E., Heine, L., Annen, J., Martial, C., Larroque, S.K., Soddu, A., Marinazzo, D.,
1452 and Laureys, S. (2018). Multifaceted brain networks reconfiguration in disorders of
1453 consciousness uncovered by co-activation patterns. Human Brain Mapping *39*, 89–103.
1454 10.1002/hbm.23826.

1455 102. Demertzi, A., Antonopoulos, G., Heine, L., Voss, H.U., Crone, J.S., De Los Angeles, C., Bahri,
1456 M.A., Di Perri, C., Vanhaudenhuyse, A., Charland-Verville, V., et al. (2015). Intrinsic functional
1457 connectivity differentiates minimally conscious from unresponsive patients. Brain *138*, 2619–
1458 2631. 10.1093/brain/awv169.

1459 103. Boveroux, P., Vanhaudenhuyse, A., and Phillips, C. (2010). Breakdown of within- and between-
1460 network Resting State during Propofol-induced Loss of Consciousness. Anesthesiology *113*,
1461 1038–1053.

1462 104. Bodien, Y.G., Threlkeld, Z.D., and Edlow, B.L. (2019). Default mode network dynamics in covert
1463 consciousness. Cortex *119*, 571–574. 10.1016/j.cortex.2019.01.014.

1464 105. Spindler, L.R.B., Luppi, A.I., Adapa, R.M., Craig, M.M., Coppola, P., Peattie, A.R.D., Manktelow,
1465 A.E., Finoia, P., Sahakian, B.J., Williams, G.B., et al. (2021). Dopaminergic brainstem
1466 disconnection is common to pharmacological and pathological consciousness perturbation.
1467 Proceedings of the National Academy of Sciences of the United States of America *118*,
1468 e2026289118. 10.1073/pnas.2026289118.

106. Huang, Z., Zhang, J., Wu, J., Mashour, G.A., and Hudetz, A.G. (2020). Temporal circuit of macroscale dynamic brain activity supports human consciousness. Science Advances *6*, 87–98. 10.1126/sciadv.aaz0087.

107. Fischer, D., Snider, S.B., Barra, M.E., Sanders, W.R., Rapalino, O., Schaefer, P., Foulkes, A.S., Bodien, Y.G., and Edlow, B.L. (2022). Disorders of Consciousness Associated With COVID-19. Neurology *98*, e315–e325. 10.1212/WNL.0000000000013067.

108. Ní Mhuircheartaigh, R., Warnaby, C., Rogers, R., Jbabdi, S., and Tracey, I. (2013). Slow-wave activity saturation and thalamocortical isolation during propofol anesthesia in humans. Science translational medicine *5*, 208ra148. 10.1126/scitranslmed.3006007.

109. Warnaby, C.E., Sleigh, J.W., Hight, D., Jbabdi, S., and Tracey, I. (2017). Investigation of Slow-wave Activity Saturation during Surgical Anesthesia Reveals a Signature of Neural Inertia in Humans. Anesthesiology *127*, 645–657. 10.1097/ALN.0000000000001759.

110. Wu, G.R., Liao, W., Stramaglia, S., Ding, J.R., Chen, H., and Marinazzo, D. (2013). A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data. Medical Image Analysis *17*, 365–374. 10.1016/j.media.2013.01.003.

111. Wu, G.R., Di Perri, C., Charland-Verville, V., Martial, C., Carrière, M., Vanhaudenhuyse, A., Laureys, S., and Marinazzo, D. (2019). Modulation of the spontaneous hemodynamic response function across levels of consciousness. NeuroImage *200*, 450–459. 10.1016/j.neuroimage.2019.07.011.

112. Faes, L., Mijatovic, G., Antonacci, Y., Pernice, R., Barà, C., Sparacino, L., Sammartino, M., Porta, A., Marinazzo, D., and Stramaglia, S. (2022). A New Framework for the Time- and Frequency-Domain Assessment of High-Order Interactions in Networks of Random Processes. IEEE Transactions on Signal Processing *70*, 5766–5777. 10.1109/TSP.2022.3221892.

113. Oizumi, M., Albantakis, L., and Tononi, G. (2014). From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0. PLoS Computational Biology *10*. 10.1371/journal.pcbi.1003588.

114. Albantakis, L., Barbosa, L., Findlay, G., Grasso, M., Haun, A.M., Marshall, W., Mayner, W.G.P., Zaeemzadeh, A., Boly, M., Juel, B.E., et al. (2023). Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms. PLoS Comput Biol *19*, e1011465. 10.1371/journal.pcbi.1011465.

115. Hoel, E.P., Albantakis, L., and Tononi, G. (2013). Quantifying causal emergence shows that macro can beat micro. Proceedings of the National Academy of Sciences of the USA *110*, 19790–19795. 10.1073/pnas.1314922110.

116. Hoel, E.P., Albantakis, L., Marshall, W., and Tononi, G. (2016). Can the macro beat the micro? Integrated information across spatiotemporal scales. Neuroscience of Consciousness, 1–13. 10.1093/nc/niw012.

117. Klein, B., and Hoel, E. (2020). The Emergence of Informative Higher Scales in Complex Networks. Complexity *2020*. 10.1155/2020/8932526.

118. Varley, T.F., and Hoel, E. (2022). Emergence as the conversion of information: a unifying theory. Phil. Trans. R. Soc. A. *380*, 20210150. 10.1098/rsta.2021.0150.

1509 119. Varley, T.F., Luppi, A.I., Pappas, I., Naci, L., Adapa, R., Owen, A.M., Menon, D.K., and
1510 Stamatakis, E.A. (2020). Consciousness & Brain Functional Complexity in Propofol Anaesthesia.
1511 Scientific Reports *10*, 1–13. 10.1038/s41598-020-57695-3.

1512 120. Leslie, Kate., Skrzypek, Hannah., Paech, M.J., Kurowski, Irina., and Whybrow, Tracey. (2007).
1513 Dreaming During Anesthesia and Anesthetic Depth in Elective Surgery Patients. Anesthesiology
1514 *106*, 33–42. 10.1097/00000542-200701000-00010.

1515 121. Huang, Z., Vlisides, P.E., Tarnal, V.C., Janke, E.L., Keefe, K.M., Collins, M.M., McKinney, A.M.,
1516 Picton, P., Harris, R.E., Mashour, G.A., et al. (2018). Brain imaging reveals covert consciousness
1517 during behavioral unresponsiveness induced by propofol. Scientific Reports *8*, 1–11.
1518 10.1038/s41598-018-31436-z.

1519 122. Luppi, A.I., Vohryzek, J., Kringelbach, M.L., Mediano, P.A.M., Craig, M.M., Adapa, R., Carhart-
1520 Harris, R.L., Roseman, L., Pappas, I., Peattie, A.R.D., et al. (2023). Distributed harmonic patterns
1521 of structure-function dependence orchestrate human consciousness. Commun Biol *6*, 1–19.
1522 10.1038/s42003-023-04474-1.

1523 123. Huang, Z., Mashour, G.A., and Hudetz, A.G. (2023). Functional geometry of the cortex encodes
1524 dimensions of consciousness. Nat Commun *14*, 72. 10.1038/s41467-022-35764-7.

1525 124. Timmermann, C., Roseman, L., Haridas, S., Rosas, F.E., Luan, L., Kettner, H., Martell, J., Erritzoe,
1526 D., Tagliazucchi, E., Pallavicini, C., et al. (2023). Human brain effects of DMT assessed via EEG-
1527 fMRI. Proceedings of the National Academy of Sciences *120*, e2218949120.
1528 10.1073/pnas.2218949120.

1529 125. Carhart-Harris, R.L. (2018). The entropic brain - revisited. Neuropharmacology *142*, 167–178.
1530 10.1016/j.neuropharm.2018.03.010.

1531 126. Carhart-Harris, R.L., Muthukumaraswamy, S., Roseman, L., Kaelen, M., Droog, W., Murphy, K.,
1532 Tagliazucchi, E., Schenberg, E.E., Nest, T., Orban, C., et al. (2016). Neural correlates of the LSD
1533 experience revealed by multimodal neuroimaging. Proceedings of the National Academy of
1534 Sciences *113*, 201518377. 10.1073/pnas.1518377113.

1535 127. Naci, L., Sinai, L., and Owen, A.M. (2017). Detecting and interpreting conscious experiences in
1536 behaviorally non-responsive patients. NeuroImage *145*, 304–313.
1537 10.1016/j.neuroimage.2015.11.059.

1538 128. Naci, L., Cusack, R., Anello, M., and Owen, A.M. (2014). A common neural code for similar
1539 conscious experiences in different individuals. Proceedings of the National Academy of
1540 Sciences of the United States of America *111*, 14277–14282. 10.1073/pnas.1407007111.

1541 129. Lee, M., Sanz, L.R.D., Barra, A., Wolff, A., Nieminen, J.O., Boly, M., Rosanova, M., Casarotto, S.,
1542 Bodart, O., Annen, J., et al. (2022). Quantifying arousal and awareness in altered states of
1543 consciousness using interpretable deep learning. Nature Communications *13*, 1–14.
1544 10.1038/s41467-022-28451-0.

1545 130. Ferrarelli, F., Massimini, M., Sarasso, S., Casali, A., Riedner, B.A., Angelini, G., Tononi, G., and
1546 Pearce, R.A. (2010). Breakdown in cortical effective connectivity during midazolam-induced
1547 loss of consciousness. Proceedings of the National Academy of Sciences *107*, 2681–2686.
1548 10.1073/pnas.0913008107.

1549 131. Bodart, O., Gosseries, O., Wannez, S., Thibaut, A., Annen, J., Boly, M., Rosanova, M., Casali,
1550 A.G., Casarotto, S., Tononi, G., et al. (2017). Measures of metabolism and complexity in the
1551 brain of patients with disorders of consciousness. Neuroimage Clin *14*, 354–362.
1552 10.1016/j.nicl.2017.02.002.

1553 132. Casali, A.G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K.R., Casarotto, S., Bruno,
1554 M.-A., Laureys, S., Tononi, G., et al. (2013). A Theoretically Based Index of Consciousness
1555 Independent of Sensory Processing and Behavior. In Science translational medicine, pp. 1–10.

1556 133. Casarotto, S., Comanducci, A., Rosanova, M., Sarasso, S., Fecchio, M., Napolitani, M., Pigorini,
1557 A., Casali, A.G., Trimarchi, P.D., Boly, M., et al. (2016). Stratification of Unresponsive Patients
1558 by an Independently Validated Index of Brain Complexity. Ann Neurol *80*, 718–729.
1559 10.1002/ana.24779.

1560 134. Sarasso, S., Boly, M., Napolitani, M., Gosseries, O., Charland-Verville, V., Casarotto, S.,
1561 Rosanova, M., Casali, A.G., Brichant, J.F., Boveroux, P., et al. (2015). Consciousness and
1562 complexity during unresponsiveness induced by propofol, xenon, and ketamine. Current
1563 Biology *25*, 3099–3105. 10.1016/j.cub.2015.10.014.

1564 135. Pal, D., Dean, J.G., Liu, T., Li, D., Watson, C.J., Hudetz, A.G., and Mashour, G.A. (2018).
1565 Differential Role of Prefrontal and Parietal Cortices in Controlling Level of Consciousness.
1566 Current Biology *28*, 2145-2152.e5. 10.1016/j.cub.2018.05.025.

1567 136. Liu, J., Lee, H.J., Weitz, A.J., Fang, Z., Lin, P., Choy, M., Fisher, R., Pinskiy, V., Tolpygo, A., Mitra,
1568 P., et al. Frequency-selective control of cortical and subcortical networks by central thalamus.
1569 eLife *4*, e09215. 10.7554/eLife.09215.

1570 137. Gent, T.C., Bandarabadi, M., Herrera, C.G., and Adamantidis, A.R. (2018). Thalamic dual control
1571 of sleep and wakefulness. Nat Neurosci *21*, 974–984. 10.1038/s41593-018-0164-7.

1572 138. Bastos, A.M., Donoghue, J.A., Brincat, S.L., Mahnke, M., Yanar, J., Correa, J., Waite, A.S.,
1573 Lundqvist, M., Roy, J., Brown, E.N., et al. (2021). Neural effects of propofol-induced
1574 unconsciousness and its reversal using thalamic stimulation. eLife *10*. 10.7554/ELIFE.60824.

1575 139. Redinbaugh, M.J., Phillips, J.M., Kambi, N.A., Mohanta, S., Andryk, S., Dooley, G.L., Afrasiabi,
1576 M., Raz, A., and Saalmann, Y.B. (2020). Thalamus modulates consciousness via layer-specific
1577 control of cortex. Neuron *106*, 66-75.e12.

1578 140. Tasserie, J., Uhrig, L., Sitt, J.D., Manasova, D., Dupont, M., Dehaene, S., and Jarraya, B. (2022).
1579 Deep brain stimulation of the thalamus restores signatures of consciousness in a nonhuman
1580 primate model. Sci. Adv *8*, 5547. 10.1126/SCIADV.ABL5547.

1581 141. Luppi, A.I., Uhrig, L., Tasserie, J., Signorelli, C.M., Stamatakis, E., Destexhe, A., Jarraya, B., and
1582 Cofre, R. (2023). Local Orchestration of Global Functional Patterns Supporting Loss and
1583 Restoration of Consciousness in the Primate Brain. Preprint at bioRxiv,
1584 10.1101/2023.06.30.547281 10.1101/2023.06.30.547281.

1585 142. Redinbaugh, M.J., Afrasiabi, M., Phillips, J.M., Kambi, N.A., Mohanta, S., Raz, A., and Saalmann,
1586 Y.B. (2022). Thalamic deep brain stimulation paradigm to reduce consciousness: Cortico-striatal
1587 dynamics implicated in mechanisms of consciousness. PLoS Comput Biol *18*, e1010294.
1588 10.1371/journal.pcbi.1010294.

1589 143. Afrasiabi, M., Redinbaugh, M.J., Phillips, J.M., Kambi, N.A., Mohanta, S., Raz, A., Haun, A.M.,
1590 and Saalmann, Y.B. (2021). Consciousness depends on integration between parietal cortex,
1591 striatum, and thalamus. Cell Syst *12*, 363-373.e11. 10.1016/j.cels.2021.02.003.

1592 144. Schiff, N.D. (2008). Central thalamic contributions to arousal regulation and neurological
1593 disorders of consciousness. Ann N Y Acad Sci *1129*, 105–118. 10.1196/annals.1417.029.

1594 145. Cain, J.A., Spivak, N.M., Coetzee, J.P., Crone, J.S., Johnson, M.A., Lutkenhoff, E.S., Real, C.,
1595 Buitrago-Blanco, M., Vespa, P.M., Schnakers, C., et al. (2021). Ultrasonic thalamic stimulation in
1596 chronic disorders of consciousness. Brain Stimul, 301–303. 10.1016/j.brs.2021.01.008.

1597 146. Cain, J.A., Spivak, N.M., Coetzee, J.P., Crone, J.S., Johnson, M.A., Lutkenhoff, E.S., Real, C.,
1598 Buitrago-Blanco, M., Vespa, P.M., Schnakers, C., et al. (2022). Ultrasonic Deep Brain
1599 Neuromodulation in Acute Disorders of Consciousness: A Proof-of-Concept. Brain sciences *12*,
1600 428. 10.3390/BRAINSCI12040428.

1601 147. Edlow, B.L., Claassen, J., Schiff, N.D., and Greer, D.M. (2020). Recovery from disorders of
1602 consciousness: mechanisms, prognosis and emerging therapies. Nature Reviews Neurology.
1603 10.1038/s41582-020-00428-x.

1604 148. Edlow, B.L., Sanz, L.R.D., Polizzotto, L., Pouratian, N., Rolston, J.D., Snider, S.B., Thibaut, A.,
1605 Stevens, R.D., and Gosseries, O. (2021). Therapies to Restore Consciousness in Patients with
1606 Severe Brain Injuries: A Gap Analysis and Future Directions. Neurocritical care *35*, 68–85.
1607 10.1007/S12028-021-01227-Y.

1608 149. Schreiber, T. (2000). Measuring information transfer. Physical Review Letters *85*, 461–464.
1609 10.1103/PhysRevLett.85.461.

1610 150. Massey, J. (1990). Causality, feedback and directed information. Proc. Int. Symp. Inf. Theory
1611 Applic.(ISITA-90), 27–30.

1612 151. Williams, P.L., and Beer, R.D. (2011). Generalized Measures of Information Transfer. arXiv.
1613 arXiv:1102.1507.

1614 152. Luppi, A.I., Rosas, F.E., Deco, G., Kringelbach, M.L., and Mediano, P.A.M. (2023). Information
1615 decomposition reveals hidden high-order contributions to temporal irreversibility. Preprint at
1616 arXiv, 10.48550/arXiv.2308.05664 10.48550/arXiv.2308.05664.

1617 153. Shine, J.M., Bissett, P.G., Bell, P.T., Koyejo, O., Balsters, J.H., Gorgolewski, K.J., Moodie, C.A.,
1618 and Poldrack, R.A. (2016). The Dynamics of Functional Brain Networks: Integrated Network
1619 States during Cognitive Task Performance. Neuron *92*, 544–554.
1620 10.1016/j.neuron.2016.09.018.

1621 154. Zamani Esfahlani, F., Jo, Y., Faskowitz, J., Byrge, L., Kennedy, D.P., Sporns, O., and Betzel, R.F.
1622 (2020). High-amplitude cofluctuations in cortical activity drive functional connectivity.
1623 Proceedings of the National Academy of Sciences of the United States of America *117*, 28393–
1624 28401. 10.1073/pnas.2005531117.

1625 155. Faskowitz, J., Zamani Esfahlani, F., Jo, Y., Sporns, O., and Betzel, R.F. Edge-centric functional
1626 network representations of human cerebral cortex reveal overlapping system-level
1627 architecture. Nature Neuroscience *23*, 1644–1654. 10.1038/s41593-020-00719-y.

156. Lurie, D.J., Kessler, D., Bassett, D.S., Betzel, R.F., Breakspear, M., Kheilholz, S., Kucyi, A., Liégeois, R., Lindquist, M.A., McIntosh, A.R., et al. (2020). Questions and controversies in the study of time-varying functional connectivity in resting fMRI. Network Neuroscience *4*, 30–69. 10.1162/netn_a_00116.

157. Cabral, J., Fernandes, F.F., and Shemesh, N. (2023). Intrinsic macroscale oscillatory modes driving long range functional connectivity in female rat brains detected by ultrafast fMRI. Nat Commun *14*, 375. 10.1038/s41467-023-36025-x.

158. Raut, R.V., Snyder, A.Z., Mitra, A., Yellin, D., Fujii, N., Malach, R., and Raichle, M.E. (2021). Global waves synchronize the brain's functional systems with fluctuating arousal. Science Advances *7*, eabf2709. 10.1126/SCIADV.ABF2709.

159. Vidaurre, D., Smith, S.M., and Woolrich, M.W. (2017). Brain network dynamics are hierarchically organized in time. Proceedings of the National Academy of Sciences of the United States of America *114*, 12827–12832. 10.1073/pnas.1705120114.

160. Atasoy, S., Deco, G., Kringelbach, M.L., and Pearson, J. (2018). Harmonic Brain Modes: A Unifying Framework for Linking Space and Time in Brain Dynamics. Neuroscientist *24*, 277–293. 10.1177/1073858417728032.

161. Demertzi, A., Tagliazucchi, E., Dehaene, S., Deco, G., Barttfeld, P., Raimondo, F., Martial, C., Fernández-Espejo, D., Rohaut, B., Voss, H.U., et al. (2019). Human consciousness is supported by dynamic complex patterns of brain signal coordination. Science Advances *5*, 1–12. 10.1126/sciadv.aat7603.

162. Gutierrez-Barragan, D., Singh, N.A., Alvino, F.G., Coletta, L., Rocchi, F., De Guzman, E., Galbusera, A., Uboldi, M., Panzeri, S., and Gozzi, A. (2021). Unique spatiotemporal fMRI dynamics in the awake mouse brain. Current biology *32*, 1–14. 10.1016/J.CUB.2021.12.015.

163. Luppi, A.I., Carhart-Harris, R.L., Roseman, L., Pappas, I., Menon, D.K., and Stamatakis, E.A. (2021). LSD alters dynamic integration and segregation in the human brain. NeuroImage *227*, 117653. 10.1016/j.neuroimage.2020.117653.

164. Luppi, A.I., Golkowski, D., Ranft, A., Ilg, R., Jordan, D., Menon, D.K., and Stamatakis, E.A. (2021). Brain network integration dynamics are associated with loss and recovery of consciousness induced by sevoflurane. Human Brain Mapping *42*, 2802–2822. 10.1002/hbm.25405.

165. Lord, L.D., Expert, P., Atasoy, S., Roseman, L., Rapuano, K., Lambiotte, R., Nutt, D.J., Deco, G., Carhart-Harris, R.L., Kringelbach, M.L., et al. (2019). Dynamical exploration of the repertoire of brain networks at rest is modulated by psilocybin. NeuroImage *199*, 127–142. 10.1016/j.neuroimage.2019.05.060.

166. Barttfeld, P., Uhrig, L., Sitt, J.D., Sigman, M., Jarraya, B., and Dehaene, S. (2015). Signature of consciousness in the dynamics of resting-state brain activity. Proceedings of the National Academy of Sciences *112*, 887–892. 10.1073/pnas.1418031112.

167. Uhrig, L., Sitt, J.D., Jacob, A., Tasserie, J., Barttfeld, P., Dupont, M., Dehaene, S., and Jarraya, B. (2018). Resting-state Dynamics as a Cortical Signature of Anesthesia in Monkeys. Anesthesiology *129*, 942–958. 10.1097/ALN.0000000000002336.

168. Stevner, A.B.A., Vidaurre, D., Cabral, J., Rapuano, K., Nielsen, S.F.V., Tagliazucchi, E., Laufs, H., Vuust, P., Deco, G., Woolrich, M.W., et al. (2019). Discovery of key whole-brain transitions and

1669 dynamics during human wakefulness and non-REM sleep. Nature Communications *10*.
1670 10.1038/s41467-019-08934-3.

1671 169. Bayne, T., Seth, A.K., and Massimini, M. (2020). Are There Islands of Awareness? Trends in
1672 Neurosciences *43*, 6–16. 10.1016/j.tins.2019.11.003.

1673 170. Connor, D., and Shanahan, M. (2007). A simulated global neuronal workspace with stochastic
1674 wiring. AI and Consciousness: Theoretical Foundations and Current Approaches, 43–48.

1675 171. Shanahan, M. (2006). A cognitive architecture that combines internal simulation with a global
1676 workspace. Consciousness and Cognition *15*, 433–449. 10.1016/j.concog.2005.11.005.

1677 172. Langdon, A., Botvinick, M., Nakahara, H., Tanaka, K., Matsumoto, M., and Kanai, R. (2022).
1678 Meta-learning, social cognition and consciousness in brains and machines. Neural Networks
1679 *145*, 80–89. 10.1016/J.NEUNET.2021.10.004.

1680 173. VanRullen, R., and Kanai, R. (2021). Deep learning and the Global Workspace Theory. Trends in
1681 Neurosciences. 10.1016/j.tins.2021.04.005.

1682 174. Proca, A.M., Rosas, F.E., Luppi, A.I., Bor, D., Crosby, M., and Mediano, P.A.M. (2022).
1683 Synergistic information supports modality integration and flexible learning in neural networks
1684 solving multiple tasks. Preprint at arXiv.

1685 175. Luppi, A.I., Mediano, P.A.M., Rosas, F.E., Allanson, J., Pickard, J.D., Williams, G.B., Craig, M.M.,
1686 Finoia, P., Peattie, A.R.D., Coppola, P., et al. (2022). Whole-brain modelling identifies distinct
1687 but convergent paths to unconsciousness in anaesthesia and disorders of consciousness.
1688 Communications biology *5*, 384. 10.1038/S42003-022-03330-Y.

1689 176. Naci, L., Haugg, A., MacDonald, A., Anello, M., Houldin, E., Naqshbandi, S., Gonzalez-Lara, L.E.,
1690 Arango, M., Harle, C., Cusack, R., et al. (2018). Functional diversity of brain networks supports
1691 consciousness and verbal intelligence. Scientific Reports *8*, 1–15. 10.1038/s41598-018-31525-z.

1692 177. Kandeepan, S., Rudas, J., Gomez, F., Stojanoski, B., Valluri, S., Owen, A.M., Naci, L., Nichols,
1693 E.S., and Soddu, A. (2020). Modeling an auditory stimulated brain under altered states of
1694 consciousness using the generalized ising model. NeuroImage *223*, 117367.
1695 10.1016/j.neuroimage.2020.117367.

1696 178. Varley, T.F., Craig, M.M., Adapa, R., Finoia, P., Williams, G., Allanson, J., Pickard, J., Menon,
1697 D.K., and Stamatakis, E.A. (2020). Fractal dimension of cortical functional connectivity
1698 networks & severity of disorders of consciousness. PLoS ONE *15*, 1–20.
1699 10.1371/journal.pone.0223812.

1700 179. Whitfield-Gabrieli, S., and Nieto-Castanon, A. (2012). Conn: A Functional Connectivity Toolbox
1701 for Correlated and Anticorrelated Brain Networks. Brain Connectivity *2*, 125–141.
1702 10.1089/brain.2012.0073.

1703 180. van Dijk, K.R.A., Sabuncu, M.R., and Buckner, R.L. (2012). The influence of head motion on
1704 intrinsic functional connectivity MRI. NeuroImage *59*, 431–438.
1705 10.1016/j.neuroimage.2011.07.044.

1706 181. Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., and Petersen, S.E. (2012). Spurious but
1707 systematic correlations in functional connectivity MRI networks arise from subject motion.
1708 NeuroImage *59*, 2142–2154. 10.1016/j.neuroimage.2011.10.018.

182. Behzadi Y, Restom K, Liau J, and Liu TT (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. NeuroImage *37*, 90–101.

183. Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., and Petersen, S.E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. NeuroImage *84*, 320–341. 10.1016/j.neuroimage.2013.08.048.

184. Lydon-Staley, D.M., Ciric, R., Satterthwaite, T.D., and Bassett, D.S. (2019). Evaluation of confound regression strategies for the mitigation of micromovement artifact in studies of dynamic resting-state functional connectivity and multilayer network modularity. Network Neuroscience *3*, 427–454. 10.1162/netn_a_00071.

185. Andellini, M., Cannatà, V., Gazzellini, S., Bernardi, B., and Napolitano, A. (2015). Test-retest reliability of graph metrics of resting state MRI functional brain networks: A review. Journal of Neuroscience Methods *253*, 183–192. 10.1016/j.jneumeth.2015.05.020.

186. Murphy, K., and Fox, M.D. (2017). Towards a consensus regarding global signal regression for resting state functional connectivity MRI. NeuroImage *154*, 169–173. 10.1016/j.neuroimage.2016.11.052.

187. Tanabe, S., Huang, Z., Zhang, J., Chen, Y., Fogel, S., Doyon, J., Wu, J., Xu, J., Zhang, J., Qin, P., et al. (2020). Altered Global Brain Signal during Physiologic, Pharmacologic, and Pathologic States of Unconsciousness in Humans and Rats. Anesthesiology, 1392–1406. 10.1097/ALN.0000000000003197.

188. Cover, T.M., and Thomas, J.A. (2005). Elements of Information Theory (Wiley-Interscience) 10.1002/047174882X.

189. Barrett, A.B. (2015). Exploration of synergistic and redundant information sharing in static and dynamical Gaussian systems. Physical Review E *91*, 52802. 10.1103/PhysRevE.91.052802.

190. Schulz, M.A., Yeo, B.T.T., Vogelstein, J.T., Mourao-Miranada, J., Kather, J.N., Kording, K., Richards, B., and Bzdok, D. (2020). Different scaling of linear models and deep learning in UK Biobank brain images vs. machine-learning datasets. Nature Communications *11*. 10.1101/757054.

191. Nozari, E., Bertolero, M.A., Stiso, J., Caciagli, L., Cornblath, E.J., He, X., Mahadevan, A.S., Pappas, G.J., and Bassett, D.S. (2023). Macroscopic resting-state brain dynamics are best described by linear models. Nat. Biomed. Eng, 1–17. 10.1038/s41551-023-01117-y.

192. Mediano, P.A.M., Rosas, F.E., Luppi, A.I., Jensen, H.J., Seth, A.K., Barrett, A.B., Carhart-Harris, R.L., and Bor, D. (2022). Greater than the parts: a review of the information decomposition approach to causal emergence. Philosophical transactions. Series A, Mathematical, physical, and engineering sciences *380*. 10.1098/RSTA.2021.0246.

193. Lizier, J.T. (2014). JIDT: An Information-Theoretic Toolkit for Studying the Dynamics of Complex Systems. Frontiers in Robotics and AI *1*, 1–37. 10.3389/frobt.2014.00011.

194. Albantakis, L., Barbosa, L., Findlay, G., Grasso, M., Haun, A.M., Marshall, W., Mayner, W.G., Zaeemzadeh, A., Boly, M., Juel, B.E., et al. (2022). Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms. Preprint at arXiv, 10.48550/arXiv.2212.14787 10.48550/arXiv.2212.14787.

195. Tegmark, M. (2016). Improved Measures of Integrated Information. PLoS Computational Biology *12*. 10.1371/journal.pcbi.1005123.

196. Barrett, A.B., and Seth, A.K. (2011). Practical Measures of Integrated Information for Time-Series Data. PLoS Comput Biol *7*, 1001052. 10.1371/journal.pcbi.1001052.

197. Oizumi, M., Tsuchiya, N., and Amari, S.I. (2016). Unified framework for information integration based on information geometry. Proceedings of the National Academy of Sciences of the United States of America *113*, 14817–14822. 10.1073/pnas.1603583113.

198. Barrett, A.B., and Mediano, P.A.M. (2019). The phi measure of integrated information is not well-defined for general physical systems. Journal of Consciousness Studies *26*, 11–20.

199. Barbosa, L.S., Marshall, W., Albantakis, L., and Tononi, G. (2021). Mechanism Integrated Information. Entropy *23*, 362. 10.3390/e23030362.

200. Rubinov, M., and Sporns, O. (2011). Weight-conserving characterization of complex functional brain networks. NeuroImage *56*, 2068–2079. 10.1016/j.neuroimage.2011.03.069.

201. Lehmann, E.L., and Romano, J.P. (2005). Testing Statistical Hypotheses Third Edition With 6 Illustrations.

202. Markello, R.D., and Misic, B. (2021). Comparing spatial null models for brain maps. NeuroImage *236*, 118052. 10.1016/j.neuroimage.2021.118052.

203. Váša, F., and Mišić, B. (2022). Null models in network neuroscience. Nat Rev Neurosci *23*, 493–504. 10.1038/s41583-022-00601-9.

204. Alexander-Bloch, A.F., Shou, H., Liu, S., Satterthwaite, T.D., Glahn, D.C., Shinohara, R.T., Vandekar, S.N., and Raznahan, A. (2018). On testing for spatial correspondence between maps of human brain structure and function. NeuroImage *178*, 540–551. 10.1016/j.neuroimage.2018.05.070.