Thermodynamics of consciousness: A non-invasive perturbational framework

Tomas Berjaga-Buisan [1*], Juan Manuel Monti [2,3], Martina Cortada [4,5], Michele A. Colombo [6], Sebastian M. Geli [1], Gianluca Gaglioti [6], Simone Sarasso [6], Morten L. Kringelbach [7,8,9,10], Maurizio Corbetta [11,12,13], Maria V. Sanchez-Vives [4,14], Marcello Massimini [6,15], Yonatan Sanz Perl [1,10,16], Gustavo Deco [7,10,14,17]

Affiliations:

[1] Center for Brain and Cognition, Computational Neuroscience Group, Department of Engineering, Universitat Pompeu Fabra, Edifici Mercè Rodoreda, Carrer Trias i Fargas 25-27, 08005, Barcelona, Catalonia, Spain.

[2] Instituto de Física Rosario CONICET-UNR, Rosario, Argentina.

[3] Laboratorio de Colisiones Atómicas, FCEIA, Universidad Nacional de Rosario, Rosario, Argentina.

[4] Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Roselló 149-153, 08036, Barcelona, Spain.

[5] Facultat de Física, Universitat de Barcelona (UB), Carrer de Martí i Franquès, 1-11, 08028 Barcelona, Spain.

[6] Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milan, Italy.

[7] Centre for Eudaimonia and Human Flourishing, Linacre College, University of Oxford, Oxford, UK.

[8] Department of Psychiatry, University of Oxford, Oxford, UK.

[9] Center for Music in the Brain, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark.

[10] International Centre for Flourishing, Universities of Oxford (UK), Aarhus (Denmark) and Pompeu Fabra (Spain).

[11] Padova Neuroscience Center (PNC), University of Padova, via Orus 2/B, 35129 Padova, Italy.

[12] Department of Neuroscience University of Padova, via Giustiniani 5, 35128 Padova, Italy.

[13] Venetian Institute of Molecular Medicine (VIMM), via Orus 2/B, 35129 Padova, Italy.

[14] Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig de Lluís Companys, 23, 08010 Barcelona, Spain.

[15] IRCCS, Fondazione Don Carlo Gnocchi Onlus, Milan 20148, Italy.

[16] Departament of Engineering, Universidad de San Andrés, Buenos Aires, Argentina.

[17] Center for Brain and Cognition, Computational Neuroscience Group, Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Edifici Mercè Rodoreda, Carrer Trias i Fargas 25-27, 08005, Barcelona, Catalonia, Spain.

[*]**Correspondence: tomas.berjaga@upf.edu**

ABSTRACT

The quest for reliable and objective measures of consciousness is critical in basic and clinical neuroscience. Across species, the Perturbational Complexity Index (PCI) has emerged as a robust empirical marker by directly perturbing the brain, yet its underlying principles of physics are not fully understood. Here, we bridge this gap by introducing a non-invasive framework based on generative whole-brain models of non-equilibrium brain dynamics. Using these models, we identified violations of the Fluctuation-Dissipation Theorem (FDT) in humans and rodents across wakefulness, anesthesia, and disorders of consciousness. Mirroring the patterns observed with PCI, we found decreased FDT violations in unresponsive disorders of consciousness and anesthesia compared to conscious conditions. This reveals a close link between PCI and non-equilibrium dynamics in spontaneous brain signals, grounding PCI in fundamental principles of physics. Overall, this framework offers new complementary, non-invasive, model-based avenues for understanding the nature of consciousness and for developing objective tools to assess its loss and recovery in health and disease. It also provides a principled foundation for discovering novel strategies to restore consciousness.

KEYWORDS

Consciousness, Disorders of Consciousness, Anesthesia, Perturbational Complexity Index, Whole-Brain Modeling, Non-equilibrium, Thermodynamics, Fluctuation-Dissipation Theorem, EEG, LFP.

INTRODUCTION

Distinguishing between conscious and unconscious brain states remains one of the most fundamental and unresolved challenges in systems and clinical neuroscience. Clinical diagnosis currently relies on a patient's ability to interact with the environment, which is typically assessed using standardized behavioral scales, such as the Coma Recovery Scale–Revised (CRS-R) [1,2]. However, the absence of behavioral signs cannot be considered definitive evidence of unconsciousness, as individuals may be unable to interact with the environment due to impairments in motor, sensory, or executive functions [3–5]. States of disconnected consciousness can occur in healthy individuals during dreaming, under general anesthesia, and in cases of severe brain injury [6–9]. Therefore, there is an urgent need for objective neurophysiological markers capable of indexing consciousness independently of behavior.

The Perturbational Complexity Index (PCI) is a generalizable non-behavioral marker of consciousness which quantifies the algorithmic complexity of spatiotemporal brain patterns in response to external perturbations, typically delivered via transcranial magnetic stimulation (TMS) in humans or intracortical electrical stimulation in animal models [10,11]. PCI captures the presence of both functional integration and differentiation, by computing the extent and information content of the chain of interactions evoked in the brain by a direct perturbation. These features are collectively referred to as brain complexity, which has been proposed as a fundamental requirement for consciousness in both theoretical models and neuroimaging studies [12–14]. PCI has demonstrated reliability in discriminating consciousness from unconsciousness [15]. Notably, previous studies reported that PCI generalized across species, scales and experimental models, consistently distinguishing between conditions in which the brain state was systematically manipulated [16,17]. Beyond this empirical robustness, PCI can be traced back to theoretical principles, and the neuronal underpinnings of its physiological and pathological alterations have been extensively studied [16–22]. However, its relationship to broader principles of physics is unclear.

Like other living systems, the brain operates far from thermodynamic equilibrium [23,24]. Neural dynamics are characterized by continuous energy dissipation, broken temporal symmetry, and directional information [24–28]. Therefore, recent advances in non-equilibrium thermodynamics for whole-brain dynamics offer a promising mechanistic framework for addressing the conceptual limitations of complexity measures [24,29]. In particular, the Fluctuation-Dissipation Theorem (FDT) states that in thermodynamic equilibrium, a system's response to external perturbations is fully determined by the structure of its spontaneous fluctuations [30]. In non-equilibrium systems, such as the brain, this balance no longer holds, and the degree to which FDT is violated can be used to quantify the system's departure from equilibrium. Importantly, this violation is a marker of non-equilibrium dynamics and has been proposed and validated as a potential signature of consciousness in previous fMRI studies [31,32].

In this study, we hypothesized and empirically tested a theoretical link between PCI and FDT violations, proposing that both reflect a common underlying property of conscious brain states: their departure from thermodynamic equilibrium. Specifically, we suggest that the brain's capacity to generate complex cortical responses to perturbations in conscious states is supported by a breakdown in the relationship between spontaneous fluctuations and evoked activity. From this perspective, high PCI values, which reflect rich and differentiated evoked responses, are predicted to coincide with strong FDT violations, indicating that the system output cannot be inferred from its intrinsic activity. Conversely, low PCI values observed in unconscious conditions are expected to reflect weaker violations of FDT, where the relationship between spontaneous and evoked activity is more preserved. Moreover, we hypothesized that FDT violations could be a promising complementary marker of consciousness, since they can be estimated from spontaneous activity alone, providing a mechanistically grounded, stimulation-free measure of consciousness.

To test these hypotheses, we analyzed neurophysiological recordings from animal models and human subjects across multiple levels of consciousness and experimental conditions. Specifically, evoked responses and spontaneous activity were recorded in mice under three different levels of anesthesia, as well as from two independent human EEG datasets. Among these two datasets, one included patients with disorders of consciousness (DoC), encompassing individuals diagnosed with unresponsive wakefulness syndrome (UWS) and minimally conscious states (MCS+, MCS−) with both vascular and traumatic etiologies. The second dataset served as a reference and included neurotypical individuals during wakefulness and under different anesthetic agents, enabling the generalization of our findings across conscious states and providing a brain-based comparison that does not rely on behavioral responsiveness. PCI was computed from cortical responses to direct perturbations, whereas FDT violations were assessed using a whole-brain model fitted to empirical recordings of spontaneous fluctuations.

The results show that FDT violations capture different levels of consciousness at a multiscale level and can discriminate among different states across all datasets. We also found a robust and systematic correlation in each species, dataset, and measurement modality: states with higher PCI values were associated with stronger FDT violations, whereas lower PCI values corresponded to smaller departures from equilibrium.

These findings provide further mechanistic foundation for PCI and other perturbation-based indices while introducing FDT violations as a complementary, perturbation-free measure for indexing consciousness. All in all, this study bridges systems and clinical neuroscience with the foundational principles of thermodynamic physics. This demonstrates that the complex dynamics associated with consciousness emerge from the intrinsic departure of the brain from equilibrium, reflecting thermodynamic asymmetries that characterize living systems. This framework offers a unifying perspective on consciousness grounded in non-equilibrium thermodynamics and opens new

possibilities for assessing consciousness in basic research and clinical practice, as well as for understanding its neural substrates in both health and disease.

RESULTS

Conscious brain states are thought to emerge from strong deviations from thermodynamic equilibrium, with a theoretical link to cortical complexity and PCI. To test this hypothesis, large-scale neurophysiological evoked and spontaneous recordings were analyzed from multiple species, modalities, and levels of consciousness. From these recordings, PCI is estimated from evoked cortical responses as a measure of complexity, whereas FDT violations are quantified from spontaneous activity to reflect the system's non-equilibrium properties (Figure 1A).

PCI values, derived from TMS-evoked or intracortical responses, were gathered from previous studies [6,15,33,34] . In these datasets, spatiotemporal patterns of cortical source activations were binarized using non-parametric permutations, and the PCI was then computed using Lempel-Ziv compression algorithm. In parallel, we quantified FDT violations using a model-based formalism that captures the mismatch between the effective response to a small theoretical perturbation and the response that would be expected if the system obeyed equilibrium (Figure 1B). In equilibrium systems, fluctuations fully determine the response function; in the brain, departures from this relationship indicate non-equilibrium dynamics.

To obtain the effective response function, we used whole-brain models. First, the directed asymmetric interactions between brain regions were estimated by fitting a multivariate Ornstein-Uhlenbeck process to the empirical data. This approach has been successfully applied to characterize large-scale brain dynamics in both EEG and fMRI studies and has proven effective in extracting biologically meaningful structure from spontaneous activity [35–37]. The model fitting was performed using a pseudo-gradient descent algorithm that simultaneously captures the zero-lag functional correlation matrix and the time-lagged covariance matrix, the latter encoding directional temporal interactions (Figure 1C). Once the model was fit, FDT violations were computed using the off-equilibrium extension provided in previous studies [32,38,39]. After both PCI and FDT metrics were estimated, we assessed their correlation across conditions, as well as their relationship with clinical scores (Figure 1A).

We applied this pipeline across three distinct datasets: (i) invasive local field potentials (LFP) recordings from mice under graded anesthesia (N = 24), (ii) scalp EEG from neurotypical human participants in wakefulness and under three anesthetic agents (N = 30), and (iii) scalp EEG from patients with disorders of consciousness (N = 17). Within each dataset, we found that FDT violations increased from unconscious to conscious states and tightly co-varied with PCI, suggesting a unified mechanism underlying both spontaneous and evoked signatures of consciousness. These findings

provide further mechanistic foundation for PCI and other perturbation-based indices while introducing FDT violations as a complementary, perturbation-free measure for indexing consciousness.

FDT Violations Track Levels of Anesthesia in Animal Models

To investigate the relationship between PCI and FDT violations and to evaluate whether they can effectively track levels of consciousness, we first analyzed electrophysiological data from animal models. We examined large-scale LFP recordings from eight mice under three distinct levels of isoflurane anesthesia (N = 24). Both spontaneous and evoked brain activity were recorded.

As illustrated in Figure S1A, spontaneous activity during anesthesia was characterized by slow oscillations [40], reflecting the coordinated dynamics of large populations of neurons transitioning between active periods, known as Up states, and silent periods, referred to as Down states [41,42]. The depth of anesthesia was associated with a slower dominant frequency and a higher occurrence of silent periods (See *Supplementary Note 1 & Supplementary Note 2)*.

PCI, derived from multi-unit activity–based evoked responses, significantly discriminated between all anesthesia conditions and consistently tracked changes in the anesthesia across levels (Spearman's $\rho = 0.71$, $p < 0.001$; Figure 2A).

In parallel, FDT violations, derived from whole-brain models fitted to the empirical spontaneous activity, also significantly distinguished all three levels of anesthesia (Figure 2B). This was supported by a robust monotonic relationship with anesthesia depth (Spearman's $\rho = 0.885$, $p < 0.001$). FDT violations were also highly correlated with the dominant frequency of the slow oscillations (see *Supplementary Note 2*).

Figure 2C shows a strong correlation between PCI and FDT violations (Pearson's $r = 0.68$, $p < 0.001$; Spearman's $\rho = 0.68$, $p < 0.001$). A linear mixed-effects model, controlling anesthesia level and incorporating a random intercept for subject, confirmed that the association remained significant ($\beta = 0.613 \pm 0.298$ s.e.m., $p = 0.037$).

FDT Violations Distinguish Wakefulness from Anesthesia in Human EEG

To generalize the findings to humans, we analyzed EEG recordings from neurotypical individuals. These data enabled us to extend the results across distinct conscious states and evaluate our findings with brain-based markers of consciousness that do not rely on behavioral responsiveness. The dataset included fifteen healthy participants recorded during wakefulness, with eyes-open or eyes-closed conditions. Participants were randomly assigned to one of three anesthetic conditions (N = 5 each for xenon, propofol, and ketamine), resulting in a total of N = 30 recordings. Participants under ketamine anesthesia reported vivid, dream-like experiences during unresponsiveness, consistent with previous studies describing ketamine-induced conscious states [6,43,44].

As illustrated in Figure S1B, spontaneous activity during anesthesia was dominated by slow oscillations, while wakefulness was characterized by prominent alpha rhythms [22,45].

PCI successfully distinguished participants who retrospectively reported any form of conscious experience from those who did not. To evaluate its performance as a binary classifier, the area under the receiver operating characteristic (AROC) curve was computed, finding that PCI achieved an AROC of 1.0, indicating perfect classification between conscious and unconscious states in this sample. A strong monotonic relationship was also observed between PCI and conditions (Spearman's $\rho = 0.867$; $p < 0.001$), indicating that PCI could index consciousness during anesthesia conditions (Figure 3A). PCI did not differentiate the ketamine condition from wakefulness, which aligns with the presence of ongoing conscious activity despite behavioral unresponsiveness.

FDT violations also distinguished participants who reported conscious experience from those who did not. This was supported by an AROC of 1.0 and a similarly strong monotonic correlation (Spearman's $\rho = 0.845$; $p < 0.001$; Figure 3B). As with PCI, FDT violations did not separate the ketamine condition from wakefulness.

FDT violations were strongly correlated with PCI (Pearson's $r = 0.85$, $p < 0.001$; Spearman's $\rho = 0.84$, $p < 0.001$; Figure 3C). A linear mixed-effects model, controlling for anesthesia state (anesthetized vs. awake) and drug type, while including subject identity as a random intercept, confirmed that the association between FDT violations and PCI remained significant ($\beta = 0.564 \pm 0.275$ s.e.m., $p < 0.001$).

FDT Violations Correlate with PCI and Clinical Scores in Disorders of Consciousness

Lastly, to evaluate the clinical utility of our findings in the context of disorders of consciousness, we analyzed EEG data from patients diagnosed with chronic DoC, including individuals classified as UWS, MCS−, and MCS+. Clinical diagnosis was based on repeated CRS-R assessments conducted at least three times within one week. EEG recordings were obtained from 17 patients, with the cohort comprising individuals with vascular (N = 13) and traumatic (N = 4) etiologies.

Consistent with previous reports, PCI reliably distinguished between behaviorally responsive and unresponsive patients (Cliff's $\delta = 1.0$, $p < 0.001$; Figure 4A). To evaluate its performance as a binary classifier, the AROC was computed. PCI reached a value of 1.0, indicating perfect classification between conscious and unconscious states in this sample.

FDT violations also significantly differentiated behaviorally responsive from unresponsive patients (Cliff's $\delta = 0.967$, $p < 0.001$; Figure 4B), offering a complementary robust marker of consciousness that does not require external perturbation. The corresponding AROC was 0.967, confirming excellent discriminative performance in this clinical setting.

In this more challenging clinical context, FDT violations remained significantly correlated with PCI (Pearson's r = 0.62, p = 0.0076; Spearman's ρ = 0.60, p = 0.0108; Figure 4C). An ANCOVA including *FDT violations* × *Condition* interaction term revealed no evidence of slope heterogeneity between groups, indicating that the relationship was conserved across conditions ($F(1,13) = 0.007$, p = 0.934; Freedman–Lane permutation p_{perm} = 0.908). Notably, FDT violations correlated more strongly with the CRS-R scores (Pearson's r = 0.707, p = 0.0015; Spearman's ρ = 0.736, $p < 0.001$; Figure 4D), than PCI did (Pearson's r = 0.56, p = 0.018; Spearman's ρ = 0.57, p = 0.0116).

DISCUSSION

Here, we show that FDT violations provide a link between PCI and fundamental principles of physics, potentially paving the way to the implementation of novel model-based markers of consciousness.

Across three different electrophysiological datasets in mice and humans, FDT violations captured differences in consciousness levels and reliably distinguished between multiple conscious states, including wakefulness, general anesthesia, ketamine-induced unresponsiveness, and DoC. Moreover, we found a robust and systematic correlation between FDT violations and PCI in all datasets, species, and modalities. Together, these results propose and establish a unified framework linking spontaneous dynamics, perturbational complexity, and fundamental thermodynamic principles.

Currently, the assessment of consciousness relies on behavioral responsiveness, most commonly assessed with the CRS-R scale in the clinical context. However, an absence of responsiveness cannot be considered as absence of consciousness [3–5]. Motor, sensory, or executive impairments may prevent behavioral interaction with the environment, even when awareness is preserved, as in healthy subjects during dreaming, under general anesthesia, or in cases of severe brain injury [7–9,46]. This limitation has motivated the search for non-behavioral markers of consciousness. Early candidates such as the P3b component, global EEG activation, and cortical oscillations showed initial promise but have proven unreliable [47–49]. In practice, they failed to consistently separate minimally conscious from unresponsive patients or to index consciousness properly across altered states [33,50–54].

The development of new quantitative complexity indices marked an important advance, inspired by theoretical neuroscience and particularly by Integrated Information Theory [13,55–57]. Empirically, brain complexity has been quantified either through analysis of spontaneous activity [58–61] or by quantifying reactivity patterns to cortical perturbations [6,10]. Among these, PCI has emerged as a robust and generalizable marker, reliably distinguishing conscious from unconscious conditions (As also seen here in Figure 2A, Figure 3A & Figure 4A) [10,11,15]. While PCI can be traced back to theoretical principles [16–22], its relationship to broader principles of physics remains unclear.

Our study directly addresses this open question by providing a novel empirical link between PCI and FDT violations across species, brain states and modalities (Figure 2C, Figure 3C, Figure 4C). Specifically, we find that states with higher PCI values, reflecting richer, more differentiated evoked

responses, are associated with stronger FDT violations, while states with lower PCI values correspond to more equilibrium-like dynamics. These findings offer empirical support for a theoretical connection between brain complexity and non-equilibrium thermodynamics, giving a proper explanation to the successful findings from previous studies through a non-equilibrium perspective [6,10,15,16,34]. This interpretation is further supported by complementary analyses of other non-equilibrium measures, including entropy production, temporal irreversibility, and coupling asymmetry, which showed parallel reductions with loss of consciousness (see *Supplementary Note 3*). In this view, both PCI and FDT violations may reflect a common underlying property: the system's departure from thermodynamic equilibrium. Mechanistically, the brain's capacity to generate complex, integrated cortical responses in conscious states is supported by a breakdown in the relationship between spontaneous fluctuations and evoked activity. This breakdown, captured as FDT violations, indicates that the system's output cannot be fully predicted from its intrinsic fluctuations alone. In contrast, unconscious states, characterized by lower PCI, exhibit weaker FDT violations, suggesting a more constrained, predictable mapping between spontaneous and evoked dynamics.

Previous studies have reported relationships between PCI and emergent physics phenomena such as criticality [62], as well as modeling links between PCI and irreversibility arising from non-equilibrium dynamics [63]. Here, we provide the first empirical evidence that complexity-based evoked measures are related to direct estimates of FDT violations, a fundamental principle of physics governing equilibrium and its breakdown. This correspondence links conscious dynamics to non-equilibrium fluctuations, strengthening the theoretical foundation of complexity-based approaches and offering a complementary thermodynamic perspective on current theories of consciousness.

Our study also demonstrated that FDT violations could constitute a promising, complementary marker of consciousness, applicable to electrophysiological recordings in both humans and animal models. FDT violations successfully distinguished three levels of anesthesia in mice LFP data (Figure 2B). In humans, FDT violations reliably indexed states of wakefulness and anesthesia induced by xenon and propofol (Figure 3B). However, similar to PCI, they did not distinguish between wakefulness and the dissociative state induced by ketamine at anesthetic concentrations [6]. This observation aligns with prior clinical studies and with participants reporting "ketamine dreams", rich inner experiences upon awakening, often unrelated to their surgical environment [6,43,44]. These results suggest that ketamine-induced conscious states preserve cortical complexity and remain far from thermodynamic equilibrium. In patients with DoC, FDT violations were able to significantly differentiate between unresponsive (UWS) and minimally conscious (MCS+/MCS−) patients (Figure 4B) and correlated with CRS-R scores (Figure 4D). Overall, FDT violations captured meaningful distinctions in brain states across all datasets and at multiple scales, introducing a complementary, model-based approach to index consciousness.

Nonetheless, future work should assess the generalization of this methodology to more complex clinical populations, such as severe post-anoxic patients, who are readily identified by TMS-EEG as

showing no significant response [15]. In such cases, refined modeling may be required to infer asymmetrical connections and reliably estimate FDT violations, as these individuals often present distinct neurophysiological signatures, including globally reduced cortical activity leading to low-voltage EEG patterns [64–67]. This framework can also be used to capture focal alterations in brain dynamics. Following brain lesions, for example, quantifying local FDT violations could yield spatially resolved maps of non-equilibrium dynamics, offering mechanistic insight into how injury reshapes integration and pointing to potential targets for restoring wake-like activity through neuromodulation [68]. Taken together, this perspective suggests that FDT violations may offer a powerful, non-invasive, and generalizable approach for mechanistic insight into a wide range of neurological conditions.

Moreover, importantly, the relationship between FDT violations and fundamental neural dynamics remains an open area of investigation. In fact, the emergence of spontaneous slow oscillations (SO), a hallmark of low-complexity brain states, is commonly observed during non-rapid eye movement sleep, in patients with UWS, even during periods of eyes open, and in other pathological states [45,53,58,69]. These oscillations reflect cortical bistability as they are characterized by alternating phases of neuronal tonic firing and neuronal silence [70]. Experimental and computational studies across humans, rodents, and cortical slice preparations converge in identifying cortical bistability and OFF-periods as key disruptors of large-scale network integration and complexity [16–19,22]. Cortical perturbations reveal that bistability causally constrains network dynamics [17–19,22]. PCI reflects this directly, increasing with rich, specific interactions and decreasing when OFF-periods yield restricted, stereotypical responses [10]. Theoretically, FDT violations provide a natural framework to study this phenomenon. They quantify the breakdown in the relationship between spontaneous fluctuations and evoked activity, thereby capturing the brain's capacity to generate complex, integrated responses—an essential feature of conscious states

It should be noted that this study has several limitations. First, electrophysiological measures can be affected by artifacts or limited electrode coverage, although standard preprocessing minimizes such effects. In DoC patients, diagnostic reliability is further challenged by behavioral fluctuations and heterogeneous signal quality; while the CRS-R vigilance protocol mitigated this, future work should examine robustness in larger and more diverse cohorts [71,72]. Time since injury may also influence EEG features, but prior studies suggest only marginal effects, and our focus was on diagnostic capacity rather than longitudinal progression [73]. Methodologically, we employed linear stochastic dynamics for analytical tractability, which may underestimate nonlinear aspects of brain activity. FDT violation values should be interpreted cautiously across species or recording modalities, given differences in spatial resolution, signal properties, and recording conditions. The limited number of subjects per condition also constrains the precision of correlation and mixed-model estimates. Despite these caveats, the consistency of results across species, brain states, and modalities underscores the robustness of the main findings, but further validation in larger and more varied populations will be

essential. It is also worth noting that while here we implemented a model-based whole-brain framework, similar approaches could be directly applied to empirical brain dynamics under linear stochastic assumptions [74]. Importantly, the model-based perspective remains advantageous as it allows for in silico manipulations, offering a path to bridge empirical data with mechanistic theory.

All in all, this study identifies FDT violations as a complementary, reliable, mechanistically grounded, and stimulation-free marker of consciousness, consistent across species, recording modalities, and scales. Most importantly, by establishing a systematic relationship between FDT violations and PCI across brain states and conditions, we demonstrate that perturbational complexity and non-equilibrium dynamics are closely related. Finally, this study bridges brain complexity and non-equilibrium dynamics, grounding consciousness in fundamental principles of physics. Together, these findings introduce a unifying framework in which spontaneous brain dynamics, perturbational complexity, and non-equilibrium thermodynamics converge to explain brain states. This perspective opens new avenues for understanding the nature of consciousness and for developing objective tools to assess its loss and recovery in health and disease. These findings also lay the conceptual foundation for future efforts to identify strategies capable of restoring consciousness.

METHODS

Mice Electrophysiological Data

Animal Subjects

Data were obtained from a previously published study [34]. Recordings were conducted in eight adult male C57BL/6J mice bred in-house at the University of Barcelona and maintained on a 12 h light/dark cycle with ad libitum access to food and water. All procedures were approved by the Ethics Committee of the Hospital Clinic of Barcelona and adhered to Spanish regulatory laws (BOE 34/11370-421, 2013) and the European Union directive 2010/63/EU. Detailed methodological procedures can be found in the original publication [34].

Surgical Procedures and Anesthesia Protocol

Anesthesia was induced via intraperitoneal injection of ketamine (75 mg/kg) and medetomidine (1.3 mg/kg) and maintained with isoflurane in pure oxygen. Atropine (0.3 mg/kg), methylprednisolone (30 mg/kg), and mannitol (0.5 g/kg) were administered subcutaneously to reduce respiratory secretions and prevent cerebral edema. Body temperature was monitored and maintained at 37°C using a thermal blanket (RWD Life Science, China).

Mice were placed in a stereotaxic frame (SR-6M, Narishige, Japan), and a craniotomy and durotomy (−3.0 mm to +3.0 mm relative to bregma and +3.0 mm lateral to midline) was performed over one hemisphere. Three anesthesia levels were defined based on isoflurane concentrations: Deep = 1.16% ± 0.08 s.e.m.; Medium = 0.34% ± 0.06 s.e.m.; and Light = 0.1–0%. The volume delivered was 0.8 L/min and the animals were breathing freely via a tracheostomy. Each level was maintained for 20–30 min, and recordings were taken during a stable slow-oscillatory regime (~10 min post-concentration change, before microarousals) [75]. Reflexes were regularly tested to confirm anesthetic depth.

Electrophysiological Recordings

Extracellular LFP activity was recorded using 32-channel multielectrode arrays (550 µm interelectrode spacing) covering the entire exposed area of the cortex [76]. Signals were amplified by 100, high-pass filtered above 5 kHz, digitized at 5 kHz, and fed into a computer via a digitizer interface (CED 1401 and Spike2 software, Cambridge Electronic Design, UK).

For each anesthesia level, 300–500 seconds of spontaneous activity was recorded, followed by an equal duration of evoked activity used for PCI calculation. For further methodological details on the stimulation setup and protocol, readers are referred to the original publication where these procedures are described in detail [34].

Human EEG Data During Wakefulness and Anesthesia

Participants

Data were obtained from a previously published dataset [77] comprising fifteen healthy participants (5 males, aged 18–28 years) recorded during wakefulness with eyes-open or eyes-closed conditions. Participants were randomly assigned to one of three anesthetic conditions (N = 5 per condition: xenon, propofol, and ketamine), resulting in a total of N = 30 recordings [73]. All participants underwent medical screening to rule out contraindications to anesthesia and provided written informed consent. The experimental protocol was approved by the Ethics Committee of the University of Liège, Belgium. This dataset has been validated and widely used in prior research [15,62,73]. For full methodological details, refer to the original publication [73].

Experimental Design and Anesthetic Procedures

All experimental procedures were performed at the Centre Hospitalier Universitaire in Liège, Belgium. Spontaneous EEG, followed by TMS-EEG recordings, was recorded before drug administration (Ramsay Scale score 2) as well as upon reaching deep unresponsiveness in three consecutive assessments (Ramsay Scale score 6) [78]. At the end of the TMS-EEG recordings, anesthesia was discontinued, participants were allowed to recover, and assessments of responsiveness were continued every 30 seconds. Participants were continuously monitored for heart rate, oxygen saturation, blood pressure, exhaled CO_2, and temperature. Mild nausea, if present, was treated with metoclopramide (2 mg).

To enable comparisons across drugs with distinct molecular targets, all anesthetic protocols aimed to reach a common behavioral endpoint: unresponsiveness to external stimuli systematically assessed using repeated Ramsey Scale administrations. Anesthetic protocols followed established procedures for propofol [79], xenon [80], and ketamine [81].

TMS–EEG Acquisition

EEG data were recorded using a 60-channel TMS-compatible amplifier (Nexstim Plc., Finland), with online filtering (0.1–350 Hz) and sampling at 1450 Hz. The reference electrode was placed on the forehead, and the impedance was kept below 5 kΩ. Two additional channels recorded electrooculographic activity. For further methodological details on the TMS–EEG acquisition, stimulation, and navigation procedures, readers are referred to the original studies where these methods are described in detail [73].

Assessment of Conscious Experience

To assess the presence or absence of conscious experience during anesthesia-induced unresponsiveness, retrospective reports were collected in all participants after awakening.

Responsiveness was considered restored when participants consistently reached a Ramsay score of 2 on three consecutive assessments. At that point, they were asked whether they recalled any mental activity during the unresponsive state. Conscious experience was defined broadly as any kind of mental activity, including thoughts, dreams, images, or emotions. Under propofol and xenon, participants either reported no memory of the period or were unable to recall any experience. In contrast, all participants who received ketamine described vivid, emotionally intense experiences with a coherent narrative structure, consistent with previous findings [6].

Human EEG Data in Disorders of Consciousness

Patients

Data were gathered from previously published studies [15,33] and included a subset of 17 patients diagnosed with prolonged or chronic DoC. Patients were clinically classified as being in UWS, MCS−, or MCS+ based on repeated CRS-R assessments conducted at least three times within one week. The cohort comprised patients with vascular (N = 13) and traumatic (N = 4) etiologies. Patients with post-anoxic etiology were not included in this study because, in severe cases, their electrophysiological signatures primarily reflect cortical integrity rather than consciousness capacity [33]. These patients often show marked suppression of alpha power within a context of overall power suppression and low-voltage EEG [64–67,82,83].

Written informed consent was obtained either from the patients (when possible) or from their legal representatives. All experimental procedures were approved by the respective local ethics committees. All patients were free of sedative medication at the time of recording. Further methodological details can be found in the original publications [15,33].

Experimental Design

During recordings, patients were seated in an upright or reclined position in a quiet room. Short spontaneous EEG recordings (median duration = 5.13 minutes, interquartile range: 4.03–6.59 minutes) were acquired either during the same experimental session of TMS-EEG recordings (N = 14) or, when not feasible, during a closely timed session within one day (N = 1) or at most within two weeks (N = 2). In all cases, the CRS-R behavioral diagnosis remained stable across sessions. Patients were continuously monitored by trained personnel to detect any signs of drowsiness. When necessary, the CRS-R vigilance protocol [2] was administered to minimize the risk of sleep intrusions and reduce the potential confounding effects of arousal fluctuations on the variables of interest.

In the DoC population, behavioral classification based on CRS-R yielded a natural dichotomy: behaviorally unresponsive patients (B−, i.e., UWS; N = 5) and behaviorally responsive patients (B+, i.e., MCS− and MCS+; N = 12).

TMS-EEG Acquisition

EEG was recorded using Ag/AgCl electrodes and TMS-compatible amplifiers (Nexstim Plc., 60 channels, N = 14; Brain Products GmbH, 64 channels, N = 3). The two systems employed highly similar electrode montages (see *Supplementary Note 4*), with a common acquisition reference placed on the forehead near Fpz. Electrode input impedance was maintained below 20 kΩ. For further methodological details on the TMS–EEG acquisition, stimulation, and navigation procedures, readers are referred to the original studies in which these methods are described in detail [15,33].

Preprocessing

Animal Electrophysiological Signals

For rodent data, distinct signal components were extracted from the same LFP recordings. For PCI computation, multiunit activity was estimated by computing the power in the 200–1500 Hz band using 5-ms windows [40,84,85]. This spectral range captures local neuronal firing and primarily reflects fast, evoked population spiking, making it well suited for quantifying spatiotemporal complexity without contamination [86]. Further methodological details can be found in the original publication [34].

In contrast, FDT violations analysis required signals that reflect the brain's spontaneous, large-scale dynamical regime. To this end, LFPs were bandpass filtered between 0.1 and 4 Hz to isolate slow cortical oscillations, which dominate anesthesia (accounting for up to 95% of spectral power) and are particularly sensitive to global state changes. [41,70]. These slow fluctuations evolve over extended timescales and display rich temporal and spatial structure, providing a natural substrate for assessing FDT violations, which quantify a system's departure from equilibrium. Importantly, this band choice was further supported by repeating the analysis using the broadband range employed for human data (see *Supplementary Note 5*).

Human EEG signals

Unlike in rodents, human recordings spanned a broader range of brain states, including wakefulness, that involve additional spectral components such as alpha and beta rhythms. For this reason, broadband signals were preserved to capture the full repertoire of state-dependent dynamics.

Spontaneous EEG signals were bandpass filtered using a third-order, zero-phase Butterworth filter (0.5–60 Hz cutoffs) with the *filtfilt* function in MATLAB, and notch-filtered to remove 50 Hz line noise and its harmonics. Data were then downsampled to reduce computational demands (from 5000 Hz to 1000 Hz for Brain Products recordings, and from 1450 Hz to 725 Hz for Nexstim recordings).

A trained neurophysiologist manually inspected each recording to exclude segments contaminated with gross artifacts. This procedure retained a median of 4.79 minutes of clean data per condition (98.74%, interquartile range: 3.58–6.12 min, 91.06–99.27%) and excluded a median of 2 electrodes per subject (IQR: 1–4). Rejected electrodes were interpolated using spherical splines, and all signals

were re-referenced to the common average. To remove residual non-neural activity, independent component analysis was performed. Components attributable to ocular, muscular, or cardiac artifacts were visually identified and excluded. A median of 33 components (IQR: 25–41.5) were retained and back-projected to reconstruct artifact-free EEG signals. Further details on recording setups and artifact removal procedures are provided in *Supplementary Note 4*.

For FDT computation, the cleaned EEG data were low-pass filtered at 40 Hz and downsampled to the corresponding Nyquist rate to minimize high-frequency muscle artifacts and reduce computational demands during whole-brain model fitting. These preprocessing steps ensured high signal fidelity and minimized non-neural confounds, preserving the information necessary for for the calculation of FDT violations.

For methodological details on the TMS-EEG preprocessing, please refer to the original studies [6,15,33]. For illustration, representative TMS–EEG traces corresponding to high and low PCI values are shown in *Supplementary Note 6*.

Perturbational Complexity Index

PCI values were obtained from previous works [6,15,33,34]. PCI quantifies the complexity of spatiotemporal brain patterns in response to external perturbations, typically delivered via TMS or intracranial stimulation. Complexity is measured using the lossless Lempel-Ziv (LZ) data compression algorithm, normalized by the source entropy [10]. To compute the LZ complexity, TMS-evoked responses are first binarized, producing a significant spatiotemporal matrix. This is achieved using a non-parametric permutation statistical procedure, as described by Casali et al. (2013) [10]. The maximum PCI value across all stimulation sites (PCImax) is typically retained as the final score. Further methodological details for PCI computation in both humans and animal models can be found in the original studies [6,15,33,34].

It should be noted that when PCI is applied to intracranial recordings and multiunit activity, particularly in animal models, it is often referred to as intracranial multiunit PCI (imPCI). This variant extends the original index to accommodate multiunit spike-based and intracortical perturbation paradigms [16,34].

Whole-Brain Modelling

The local dynamics of each brain region were modeled as a multivariate Ornstein–Uhlenbeck (OU) process, which corresponds to the linearization of nonlinear population models such as Wilson–Cowan neurons [88]. This formulation provides a linear approximation to the stochastic dynamics of the system near a stable fixed point [89,90]. OU processes are particularly well-suited for capturing non-equilibrium dynamics, as they enable a direct and quantitative assessment of the relationship between spontaneous fluctuations and linear response functions, allowing deviations from equilibrium to be

computed [35,38,91]. Moreover, OU-based models have been successfully applied to characterize large-scale brain dynamics in both EEG and fMRI studies, and have proven effective in extracting biologically meaningful structure from spontaneous activity [35–37].

In this framework, brain activity is modeled as a network of interacting regions governed by coupled stochastic differential equations:

$$\frac{dx}{dt} = -Bx(t) + \eta(t) \tag{1}$$

where B is the drift matrix and corresponds to the asymmetric connectivity matrix inferred from empirical data. The $\eta(t)$ term corresponds to zero-mean additive white Gaussian noise. The noise is characterized by its covariance structure, $< \eta(t)\eta(t') > = 2\,\sigma^2\delta(t-t')$, where σ^2 is, by definition, symmetric and positive semidefinite.

To ensure a well-defined and stable stationary solution, we require that all eigenvalues of the drift matrix B have strictly positive real parts [90]. This condition guarantees that the associated system dynamics fulfils the Hurwitz criterion, meaning the system naturally decays and converges to a unique steady state. This requirement is essential not only for mathematical stability and well-posedness, but also for the thermodynamic and biological interpretability of the model [90]. The model describes multivariate neural activity signals while preserving the temporal structure and directional interactions required to quantify non-equilibrium dynamics and assess how departures from equilibrium relate to changes in consciousness. This approach is consistent with the methodology and previous applications of the model [36,92–94].

Model Optimization

To quantify FDT violations, we applied the multivariate OU modeling framework described in the section above. The empirical observables used for model fitting included the functional connectivity matrix ($FC^{empirical}$), reflecting zero-lag statistical dependencies across brain regions, and the normalized time-lagged covariance matrix ($FS^{empirical}(\tau)$), which captures directional temporal interactions at a fixed lag while reducing zero-lag dependencies and volume-conduction effects. These normalized time-lagged covariance matrices are generated by taking the shifted covariance matrix $KS^{empirical}(\tau)$ and normalizing each pair (i,j) by dividing $\sqrt{KS_{ii}^{empirical}(0)KS_{jj}^{empirical}(0)}$. This normalization accounts for variance across regions and emphasizes relative lagged dependencies. Importantly, incorporating time-lagged covariances allows the estimated coupling matrix B to become asymmetric, enabling the emergence of non-equilibrium dynamics and quantifiable violations of the

FDT [95,96]. The time lag τ was selected to ensure theoretical consistency with the OU formulation and to preserve sensitivity to temporal asymmetries, following previous studies [35,37].

To match these empirical observables, the model was iteratively adjusted using a heuristic pseudo-gradient descent algorithm that minimized the difference between simulated and empirical FC and $FS(\tau)$ matrices [37,97]. We updated B until the fit was fully optimized and converged to a stable solution. More specifically, the updating uses the following rule:

$$B_{ij} = B_{ij} + \alpha \left(FC_{ij}^{empirical} - FC_{ij}^{model} \right) + \delta \left(FS_{ij}^{empirical}(\tau) - FS_{ij}^{model}(\tau) \right) \tag{2}$$

The multivariate OU model allows analytic computation of both stationary and time-lagged covariances. The stationary covariance matrix was defined as $K = \langle xx^T \rangle$ and the noise covariance as Q, assumed to be symmetric and positive semidefinite. Using Itô's calculus and keeping terms to first order in dt, the evolution of K is given by the following equation [31,32,96]:

$$\frac{dK}{dt} = -BK - KB^T + Q \tag{3}$$

At the steady state ($\frac{dK}{dt} = 0$), we obtain the continuous-time Lyapunov equation:

$$BK + KB^T = Q \tag{4}$$

Solving this equation (e.g., via Lyapunov or Sylvester solvers) provides the model-derived covariance matrix K, from which the simulated functional connectivity (FC^{model}), is computed. The time-lagged covariance was then calculated analytically as:

$$KS^{model}(\tau) = e^{-\tau B} K \tag{5}$$

and normalized in the same manner as the empirical data. Note that $K(0) = K$.

B initialized as a random symmetric matrix constrained to satisfy the Hurwitz stability criterion. This approach ensured mathematical and thermodynamic stability without imposing anatomical priors. Learning rates were fixed at $\alpha = \delta = 0.00001$, and the optimization was repeated until the solution stabilized. The full optimization process was repeated 1000 times, and the results were then averaged to ensure convergence to a robust solution. We refer to the final optimized matrix as the generative effective connectivity matrix, which captures both the directional and non-equilibrium aspects of spontaneous brain dynamics [98]. Noise was assumed to be homogeneous and therefore did not need to be fitted, as demonstrated in *Supplementary Note 7*.

FDT Violations

To assess the departure from equilibrium, we adapted a model-based formalism grounded in the off-equilibrium extension of FDT [38,39,99]. In thermodynamic equilibrium, the FDT states that the system's linear response to small perturbations is entirely determined by spontaneous fluctuations. However, in systems that are out of equilibrium, such as the brain, this relationship is broken. The degree of FDT violation therefore provides a principled and physically interpretable measure of non-equilibrium dynamics [31,32].

More precisely, in equilibrium, the linear response function $R(t,t')$ of the system's state $x(t)$ to a small enough perturbation $\epsilon(t')$ is defined as [30]:

$$x(t) = \int_{t'}^{t} R(t,t')\epsilon(t')dt' \tag{6}$$

According to the FDT, this response function is directly related to the autocorrelation function of spontaneous fluctuations, $C(t,t') = \langle x(t)x(t')\rangle$, through the expression:

$$R(t,t') = \beta \frac{\partial C(t,t')}{\partial t'} \tag{7}$$

where $t' <= t$ and $\beta = 1/T$, T being the temperature of the thermal bath. In the case of non-equilibrium systems, this equality no longer holds. Instead, the linear response can be determined through generalized expressions derived from stochastic thermodynamics such as [39]:

$$R(t,t') = \frac{1}{2T}\langle x(t)\eta(t')\rangle \tag{8}$$

Based on this, two metrics that capture the degree of FDT violation have been introduced in prior studies [32,38]: the differential violation $V(t,t')$ and the integral violation $I(t,t')$.

$$V(t,t') = \frac{\partial C(t,t')}{\partial t'} - TR(t,t') \tag{9}$$

$$I(t,t') = \int_{t'}^{t} V(t,s)ds = C(t,t) - C(t,t') - T\,\chi(\mathrm{t},\mathrm{t}') \tag{10}$$

where $\chi(t,t') = \int_{t'}^{t} R(t,s)ds$ is the integrated response function, or dynamic susceptibility.

To obtain scale-free quantities that allow comparisons across datasets recorded in the same species with the same recording modality, we normalized $V(t,t')$ and $I(t,t')$ by the instantaneous autocorrelation, consistent with previous thermodynamic metrics [31,100].

$$\tilde{V}(t,t') = \frac{V(t,t')}{C(t,t)} \tag{11}$$

$$\tilde{I}(t,t') = \frac{I(t,t')}{C(t,t)} \tag{12}$$

This normalization choice is justified by strict scale invariance. Normalizing by the instantaneous autocorrelation, rather than by a fixed $C(0,0)$ or by α^2 alone, further accounts for node-wise heterogeneity and slow variance drifts, while remaining robust to preprocessing steps that rescale amplitudes (e.g., z-scoring within runs).

Alternative estimators of FDT violations exist and are theoretically equivalent given infinite data. However, finite-sample effects can lead to variability across formulations. Following previous literature, we report the most stable estimates using the definitions of $\tilde{I}$ as these are less prone to numerical instability [32].

To obtain node-level metrics of non-equilibrium, we computed the time-integrated absolute values of the violations:

$$dFDT\ violation = \sum_{t'} \int \left|\tilde{V}_j(t,t')\right| dt \tag{13}$$

$$IFDT\ violation = \sum_{t'} \int \left|\tilde{I}_j(t,t')\right| dt \tag{14}$$

Then, for each subject, we take the average over all nodes as a more global indicator of distance from equilibrium. In the present study, we focus on these global averages. However, in conditions involving more spatially localized alterations, such as neurodegenerative or focal disorders, node-specific patterns may provide additional insight [22].

Statistical Analysis

Group comparisons were conducted using non-parametric tests. For paired conditions, the Wilcoxon signed-rank test was used, followed by Holm–Bonferroni post hoc corrections. For unpaired groups, the Mann–Whitney U test was applied, and effect sizes were quantified using Cliff's δ, a distribution-

free, non-parametric measure directly interpretable as the probability of superiority. Binary classification performance was evaluated using the area under the curve (AROC) as a measure of discriminative accuracy.

To assess monotonic relationships between variables, Spearman's correlation coefficients (ρ) were calculated. Pearson's correlation coefficients (r) were also computed when relationships were assumed to be linear and variables approximately normally distributed. We tested whether the relationship between FDT violations and PCI differed across conditions using an ANCOVA that included the *FDT violations* × *Condition* interaction term. The significance of this interaction was further evaluated using a Freedman–Lane permutation test (10,000 permutations, within-condition shuffling) to obtain exact p under small-sample conditions.

In paired datasets, linear mixed-effects models (LME) were fitted to confirm the robustness of the associations between FDT violations and PCI. In mice, the LME controlled for anesthesia level (deep vs mid vs light), while in humans it controlled for anesthesia state (anesthetized vs. awake) and drug type (ketamine vs propofol vs xenon). In both cases, subject identity was included as a random intercept to account for repeated measurements.

Variables were standardized prior to regression to allow interpretation of effect sizes in standard deviation units. All p-values reported are two-tailed, and significance was defined at $p < 0.05$.

DATA AVAILABILITY

Raw LFP recordings from eight adult male C57BL/6J mice across different anesthesia levels are publicly available at the EBRAINS repository (10.25493/WKA8-Q4T). Raw EEG recordings from neurotypical participants during wakefulness and anesthesia are publicly available at Zenodo (https://doi.org/10.5281/zenodo.806176). Additional data, including raw EEG signals from patients with disorders of consciousness, along with associated metadata, are available from the authors upon reasonable request and subject to a data sharing agreement.

CODE AVAILABILITY

The code used for data analysis, modeling of spontaneous fluctuations, fitting the generative effective connectivity, computation of FDT violations, and other thermodynamic metrics is available at https://github.com/tomasberjagabuisan/Thermo-Consciousness

ACKNOWLEDGMENTS

We are grateful to all the patients and volunteers who participated in the experiments and made this work possible. We thank the members of the various laboratories involved in data acquisition for their invaluable contributions. We would like to thank Irene Acero and Paula Garcia for their insightful comments on the draft. We want to also thank Inés Blázquez for assisting with the figures.

J.M.M. acknowledges financial support from CONICET through an external scholarship for researchers. M.L.K. is supported by the Centre for Eudaimonia and Human Flourishing (funded by the Pettit and Carlsberg Foundations) and Center for Music in the Brain (funded by the Danish National Research Foundation, DNRF117). M. Corbetta is supported by Fondazione Cassa di Risparmio di Padova e Rovigo (CARIPARO) (GA 55403), the Italian Ministry of Health (NEUROCONN RF-2008-12366899, EYEMOVINSTROKE RF-2019-12369300), and the EU H2020 project euSNN, H2020-SC5–2019-2, (Grant Agreement 869505). M.V.S.V. is supported by PID2023-152918OB-I00 funded by MICIU / AEI / 10.13039/501100011033/FEDER, UE and by AGAUR 2021-SGR-01165. M. Corbetta, M.V.S.V., M.M., Y.S.P. and G.D. are supported by the project Neurological Mechanisms of Injury, and Sleep-like cellular dynamics (NEMESIS; ref. 101071900) funded by the EU ERC Synergy Horizon Europe. G.D. also acknowledges support from grant PID2022-136216NB-I00 funded by MICIU/AEI/10.13039/501100011033; by ERDF A way of making Europe, ERDF, EU; and AGAUR research support grant (2021 SGR 00917) funded by the Department of Research and Universities of the Generalitat of Catalunya.

AUTHOR CONTRIBUTION

T.B.B., J.M.M., Y.S.P., and G.D. conceptualized and designed the study. T.B.B. and J.M.M. performed the formal analysis and developed the modeling framework. T.B.B., J.M.M., M. Cortada, M.A.C., M.V.S.V., M.M., Y.S.P. and G.D. contributed to the methodology. J.M.M, M. Cortada and S.M.G. performed complementary analysis. M. Cortada, M.A.C., M.V.S.V. and M.M. provided resources. T.B.B. performed additional data curation and filtering. T.B.B. generated all visualizations with input from M.L.K. G.D. supervised the research. T.B.B. wrote the original manuscript. All authors reviewed and edited the manuscript.

DECLARATION OF INTEREST

Marcello Massimini is a co-founder and shareholder of Intrinsic Powers, a spin-off of the University of Milan. Simone Sarasso is advisor of Intrinsic Powers. This affiliation in no way affects the content of this article.

SUPPLEMENTAL INFORMATION

Document S1. Figures S1–S8, additional methodological details and illustrations, and supplementary references.

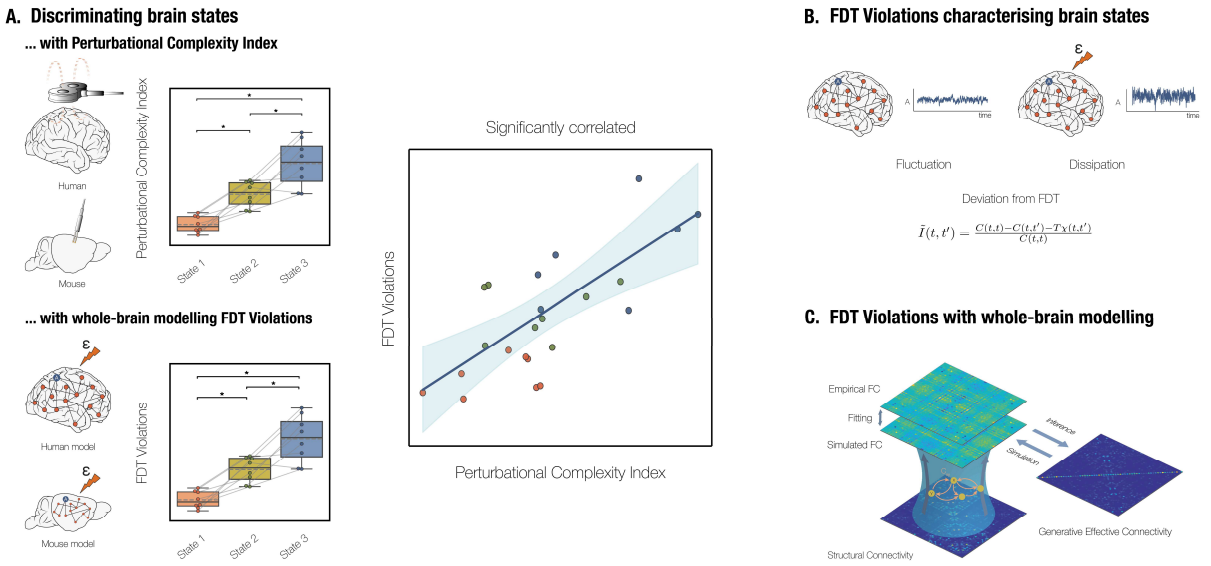

Figure 1. **Framework for computing PCI and FDT violations from evoked and spontaneous brain dynamics. (A)** PCI is computed from TMS-evoked or intracortical responses by binarizing spatiotemporal patterns of cortical source activations and quantifying their algorithmic complexity using Lempel-Ziv compression. PCI has proven to be a reliable and robust indicator, consistently distinguishing between conscious and unconscious conditions with high accuracy [10,15]. FDT violations, estimated from spontaneous brain dynamics using whole-brain modelling, also discriminate between brain states and correlate strongly with PCI. This provides a mechanistic grounding for PCI and related perturbation-based indices, while also introducing FDT violation as a complementary, perturbation-free marker of consciousness. **(B)** Here, we use a model-based formalism for FDT violations that captures the mismatch between the model-derived response to a small theoretical perturbation and the response that would be expected if the system obeyed equilibrium. In equilibrium systems, fluctuations fully determine the response function; in the brain, departures from this relationship indicate non-equilibrium dynamics. FDT violations have been used in previous studies to characterize brain states [31,32]. **(C)** Whole-brain modeling is applied at the individual level to reproduce spontaneous brain dynamics and capture the underlying asymmetrical structure of connectivity, referred to as Generative Effective Connectivity.

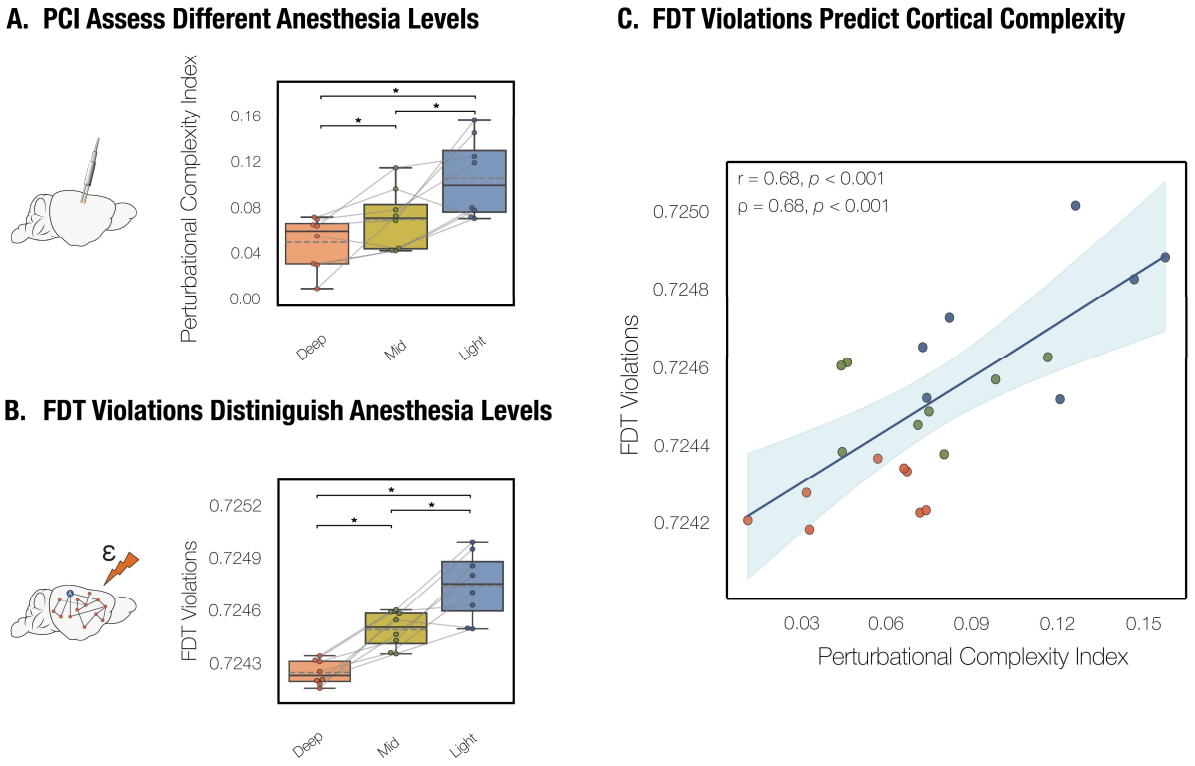

Figure 2. **FDT violations track anesthesia levels in mice and correlate with the PCI. (A)** PCI, calculated from multi-unit activity–based evoked responses, significantly discriminated between all anesthesia conditions. PCI consistently tracked the anesthesia levels in mice (Spearman's $\rho = 0.71$, $p < 0.001$). **(B)** FDT violations calculated from the whole-brain model fitted to empirical spontaneous activity were also able to discriminate among all the conditions, as reflected by a strong monotonic relationship (Spearman's $\rho = 0.885$, $p < 0.001$). **(C)** FDT violations correlate with the PCI (Pearson's $r = 0.68$, $p < 0.001$; Spearman's $\rho = 0.68$, $p < 0.001$). A linear mixed-effects model, controlling for anesthesia level and incorporating a random intercept for subject, confirmed that the association between remained significant ($\beta = 0.613 \pm 0.298$ s.e.m., $p = 0.037$). All the pairwise comparisons were performed using Wilcoxon signed-rank tests and post hoc corrected with Holm-Bonferroni. * $p < 0.05$

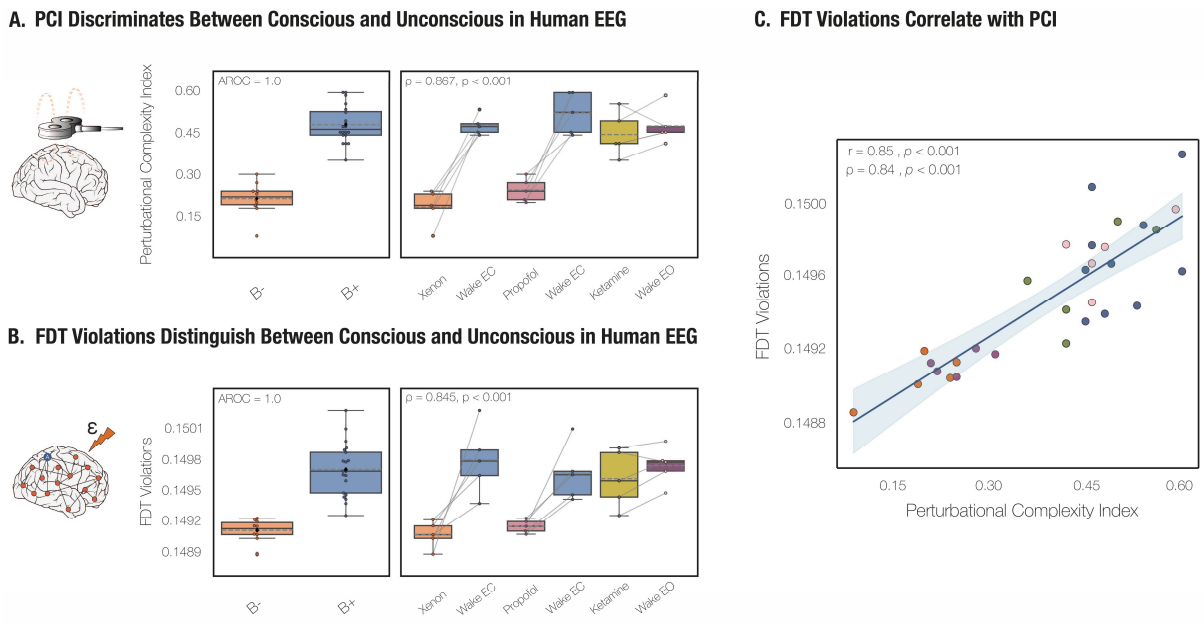

Figure 3. **FDT violations in human EEG track level of consciousness during wakefulness and anesthesia and correlate with PCI. (A)** PCI discriminated between participants who retrospectively reported some form of conscious experience (B+) and those who did not (B-). PCI achieved a perfect classification performance (AROC = 1.0). A strong monotonic relationship with condition was observed (Spearman's $\rho = 0.867$, $p < 0.001$). **(B)** FDT violations, computed from whole-brain models fitted to spontaneous EEG activity, also differentiated participants who reported conscious experience from those who did not (AROC = 1.0). A monotonic relationship was observed (Spearman's $\rho = 0.845$, $p < 0.001$), consistent with sensitivity to differences in conscious state. **(C)** FDT violations were strongly correlated with PCI (Pearson's $r = 0.85$, $p < 0.001$; Spearman's $\rho = 0.84$, $p < 0.001$). A linear mixed-effects model, controlling for anesthesia state and drug type while including subject identity as a random intercept, confirmed that the association between FDT violations and PCI remained significant ($\beta = 0.564 \pm 0.275$ s.e.m., $p < 0.001$).

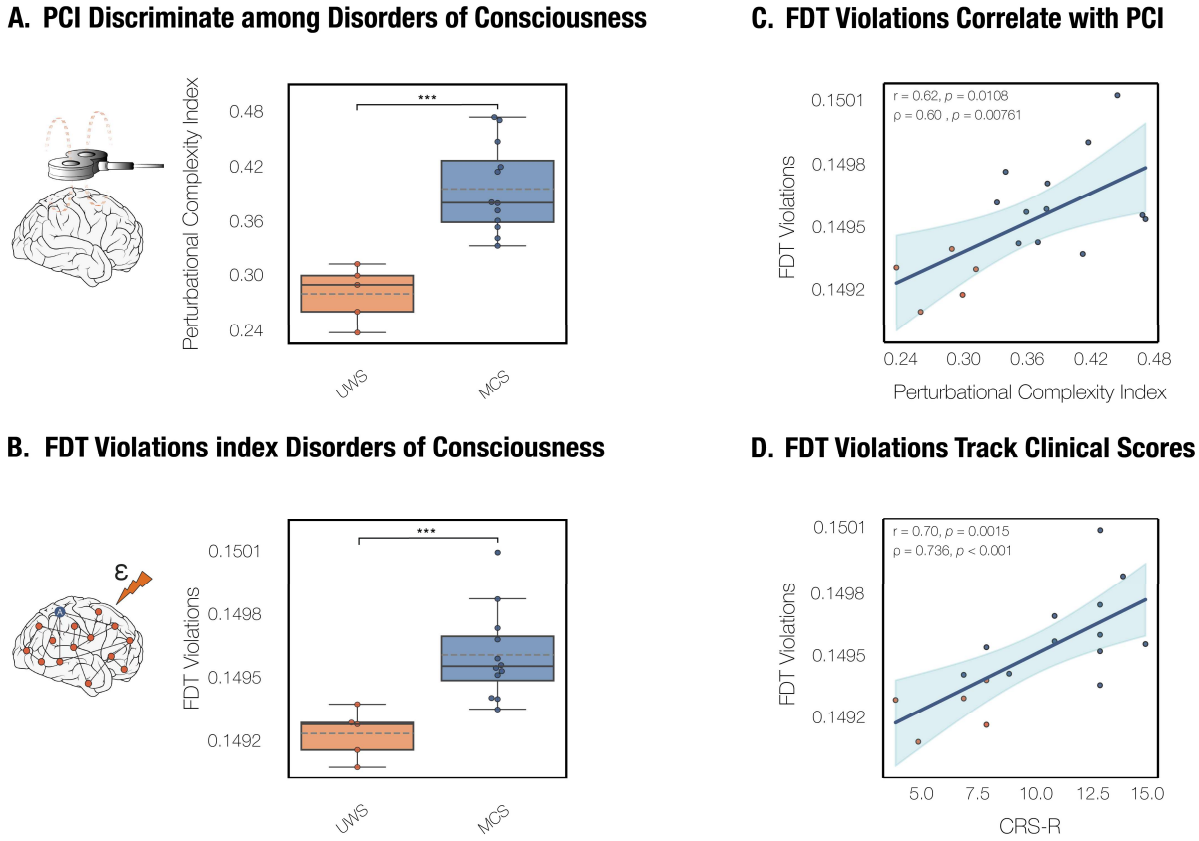

Figure 4. **FDT violations distinguish conscious from unconscious states in patients with disorders of consciousness and correlate with PCI and clinical scores. (A)** PCI discriminated between conscious (MCS-/MCS+) and unconscious patients (UWS) with high accuracy (Cliff's $\delta = 1.0$, $p < 0.001$) and achieved perfect classification performance (AROC = 1.0). **(B)** FDT violations also distinguished conscious (MCS-/MCS+) from unconscious (UWS) patients with similarly high accuracy (Cliff's $\delta = 0.967$, $p < 0.001$) and achieved an AROC of 0.967. **(C)** FDT violations were significantly correlated with PCI values, indicating that both metrics capture related aspects of conscious in a clinical context (Pearson's r = 0.62, $p = 0.0108$; Spearman's ρ = 0.60, $p = 0.0076$). An ANCOVA including the *FDT violations* × *Condition* interaction term revealed no evidence of slope heterogeneity between groups, indicating that the relationship was conserved across conditions ($F(1,13) = 0.007$, $p = 0.934$; Freedman–Lane permutation $p_{perm} = 0.908$). **(D)** FDT violations were also strongly correlated to clinical scales such as CRS-R scores (Pearson's r = 0.707, $p = 0.0015$; Spearman's ρ = 0.736, $p < 0.001$). All p-values were calculated using two-sided Mann-Whitney U test as the groups were independent. *** $p < 0.001$.

REFERENCES

1. Giacino, J.T., Schnakers, C., Rodriguez-Moreno, D., Kalmar, K., Schiff, N., and Hirsch, J. (2009). Behavioral assessment in patients with disorders of consciousness: gold standard or fool's gold? Prog. Brain Res., 33–48. https://doi.org/10.1016/s0079-6123(09)17704-x.

2. Kalmar, K., and Giacino, J.T. (2005). The JFK coma recovery scale—revised. Neuropsychol. Rehabil. *15*, 454–460. https://doi.org/10.1080/09602010443000425.

3. Alkire, M.T., Hudetz, A.G., and Tononi, G. (2008). Consciousness and Anesthesia. Science *322*, 876–880. https://doi.org/10.1126/science.1149213.

4. Laureys, S., Owen, A.M., and Schiff, N.D. (2004). Brain function in coma, vegetative state, and related disorders. Lancet Neurol. *3*, 537–546. https://doi.org/10.1016/s1474-4422(04)00852-x.

5. Sanders, R.D., Tononi, G., Laureys, S., Sleigh, J.W., and Warner, D.S. (2012). Unresponsiveness ≠ Unconsciousness. Anesthesiology *116*, 946–959. https://doi.org/10.1097/aln.0b013e318249d0a7.

6. Sarasso, S., Boly, M., Napolitani, M., Gosseries, O., Charland-Verville, V., Casarotto, S., Rosanova, M., Casali, A.G., Brichant, J.-F., Boveroux, P., et al. (2015). Consciousness and Complexity during Unresponsiveness Induced by Propofol, Xenon, and Ketamine. Curr. Biol. *25*, 3099–3105. https://doi.org/10.1016/j.cub.2015.10.014.

7. Sebel, P.S., Bowdle, T.A., Ghoneim, M.M., Rampil, I.J., Padilla, R.E., Gan, T.J., and Domino, K.B. (2004). The Incidence of Awareness During Anesthesia: A Multicenter United States Study. Anesth. Analg. *99*, 833–839. https://doi.org/10.1213/01.ane.0000130261.90896.6c.

8. Laureys, S., Celesia, G.G., Cohadon, F., Lavrijsen, J., León-Carrión, J., Sannita, W.G., Sazbon, L., Schmutzhard, E., Von Wild, K.R., Zeman, A., et al. (2010). Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. BMC Med. *8*. https://doi.org/10.1186/1741-7015-8-68.

9. Fernández-Espejo, D., and Owen, A.M. (2013). Detecting awareness after severe brain injury. Nat. Rev. Neurosci. *14*, 801–809. https://doi.org/10.1038/nrn3608.

10. Casali, A.G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K.R., Casarotto, S., Bruno, M.-A., Laureys, S., Tononi, G., et al. (2013). A Theoretically Based Index of Consciousness Independent of Sensory Processing and Behavior. Sci. Transl. Med. *5*. https://doi.org/10.1126/scitranslmed.3006294.

11. Comolatti, R., Pigorini, A., Casarotto, S., Fecchio, M., Faria, G., Sarasso, S., Rosanova, M., Gosseries, O., Boly, M., Bodart, O., et al. (2019). A fast and general method to empirically estimate the complexity of brain responses to transcranial and intracranial stimulations. Brain Stimulat. *12*, 1280–1289. https://doi.org/10.1016/j.brs.2019.05.013.

12. Tononi, G. (1998). Consciousness and Complexity. Science *282*, 1846–1851. https://doi.org/10.1126/science.282.5395.1846.

13. Deco, G., Tononi, G., Boly, M., and Kringelbach, M.L. (2015). Rethinking segregation and integration: contributions of whole-brain modelling. Nat. Rev. Neurosci. *16*, 430–439. https://doi.org/10.1038/nrn3963.

14. Sarasso, S., Casali, A.G., Casarotto, S., Rosanova, M., Sinigaglia, C., and Massimini, M. (2021). Consciousness and complexity: a consilience of evidence. Neurosci. Conscious. *2021*, niab023. https://doi.org/10.1093/nc/niab023.

15. Casarotto, S., Comanducci, A., Rosanova, M., Sarasso, S., Fecchio, M., Napolitani, M., Pigorini, A., G. Casali, A., Trimarchi, P.D., Boly, M., et al. (2016). Stratification of unresponsive patients by an independently validated index of brain complexity. Ann. Neurol. *80*, 718–729. https://doi.org/10.1002/ana.24779.

16. D'Andola, M., Rebollo, B., Casali, A.G., Weinert, J.F., Pigorini, A., Villa, R., Massimini, M., and Sanchez-Vives, M.V. (2018). Bistability, Causality, and Complexity in Cortical Networks: An In Vitro Perturbational Study. Cereb. Cortex *28*, 2233–2242. https://doi.org/10.1093/cercor/bhx122.

17. Cavelli, M.L., Mao, R., Findlay, G., Driessen, K., Bugnon, T., Tononi, G., and Cirelli, C. (2023). Sleep/wake changes in perturbational complexity in rats and mice. iScience *26*, 106186. https://doi.org/10.1016/j.isci.2023.106186.

18. Pigorini, A., Sarasso, S., Proserpio, P., Szymanski, C., Arnulfo, G., Casarotto, S., Fecchio, M., Rosanova, M., Mariotti, M., Lo Russo, G., et al. (2015). Bistability breaks-off deterministic responses to intracortical stimulation during non-REM sleep. NeuroImage *112*, 105–113. https://doi.org/10.1016/j.neuroimage.2015.02.056.

19. Rosanova, M., Fecchio, M., Casarotto, S., Sarasso, S., Casali, A.G., Pigorini, A., Comanducci, A., Seregni, F., Devalle, G., Citerio, G., et al. (2018). Sleep-like cortical OFF-periods disrupt causality and complexity in the brain of unresponsive wakefulness syndrome patients. Nat. Commun. *9*, 4427. https://doi.org/10.1038/s41467-018-06871-1.

20. Arena, A., Comolatti, R., Thon, S., Casali, A.G., and Storm, J.F. (2021). General Anesthesia Disrupts Complex Cortical Dynamics in Response to Intracranial Electrical Stimulation in Rats. eneuro *8*, ENEURO.0343-20.2021. https://doi.org/10.1523/ENEURO.0343-20.2021.

21. Cattani, A., Galluzzi, A., Fecchio, M., Pigorini, A., Mattia, M., and Massimini, M. (2023). Adaptation Shapes Local Cortical Reactivity: From Bifurcation Diagram and Simulations to Human Physiological and Pathological Responses. eneuro *10*, ENEURO.0435-22.2023. https://doi.org/10.1523/ENEURO.0435-22.2023.

22. Massimini, M., Corbetta, M., Sanchez-Vives, M.V., Andrillon, T., Deco, G., Rosanova, M., and Sarasso, S. (2024). Sleep-like cortical dynamics during wakefulness and their network effects following brain injury. Nat. Commun. *15*, 7207. https://doi.org/10.1038/s41467-024-51586-1.

23. Nartallo-Kaluarachchi, R., Asllani, M., Deco, G., Kringelbach, M.L., Goriely, A., and Lambiotte, R. (2024). Broken detailed balance and entropy production in directed networks. Phys. Rev. E *110*, 034313. https://doi.org/10.1103/PhysRevE.110.034313.

24. Kringelbach, M.L., Sanz Perl, Y., and Deco, G. (2024). The Thermodynamics of Mind. Trends Cogn. Sci. *28*, 568–581. https://doi.org/10.1016/j.tics.2024.03.009.

25. Sanz Perl, Y., Bocaccio, H., Pallavicini, C., Pérez-Ipiña, I., Laureys, S., Laufs, H., Kringelbach, M., Deco, G., and Tagliazucchi, E. (2021). Nonequilibrium brain dynamics as a signature of consciousness. Phys. Rev. E *104*. https://doi.org/10.1103/physreve.104.014411.

26. Deco, G., Sanz Perl, Y., De La Fuente, L., Sitt, J.D., Yeo, B.T.T., Tagliazucchi, E., and Kringelbach, M.L. (2023). The arrow of time of brain signals in cognition: Potential intriguing role of parts of the default mode network. Netw. Neurosci. *7*, 966–998. https://doi.org/10.1162/netn_a_00300.

27. Deco, G., Sanz Perl, Y., Bocaccio, H., Tagliazucchi, E., and Kringelbach, M.L. (2022). The INSIDEOUT framework provides precise signatures of the balance of intrinsic and extrinsic dynamics in brain states. Commun. Biol. *5*. https://doi.org/10.1038/s42003-022-03505-7.

28. Camassa, A., Torao-Angosto, M., Manasanch, A., Kringelbach, M.L., Deco, G., and Sanchez-Vives, M.V. (2024). The temporal asymmetry of cortical dynamics as a signature of brain states. Sci. Rep. *14*, 24271. https://doi.org/10.1038/s41598-024-74649-1.

29. Lynn, C.W., Cornblath, E.J., Papadopoulos, L., Bertolero, M.A., and Bassett, D.S. (2021). Broken detailed balance and entropy production in the human brain. Proc. Natl. Acad. Sci. *118*. https://doi.org/10.1073/pnas.2109889118.

30. Kubo, R. (1957). Statistical-Mechanical Theory of Irreversible Processes. I. General Theory and Simple Applications to Magnetic and Conduction Problems. J. Phys. Soc. Jpn. *12*, 570–586. https://doi.org/10.1143/JPSJ.12.570.

31. Deco, G., Lynn, C.W., Sanz Perl, Y., and Kringelbach, M.L. (2023). Violations of the fluctuation-dissipation theorem reveal distinct nonequilibrium dynamics of brain states. Phys. Rev. E *108*, 064410. https://doi.org/10.1103/PhysRevE.108.064410.

32. Monti, J.M., Perl, Y.S., Tagliazucchi, E., Kringelbach, M.L., and Deco, G. (2025). Fluctuation-dissipation theorem and the discovery of distinctive off-equilibrium signatures of brain states. Phys. Rev. Res. *7*, 013301. https://doi.org/10.1103/PhysRevResearch.7.013301.

33. Colombo, M.A., Comanducci, A., Casarotto, S., Derchi, C.-C., Annen, J., Viganò, A., Mazza, A., Trimarchi, P.D., Boly, M., Fecchio, M., et al. (2023). Beyond alpha power: EEG spatial and spectral gradients robustly stratify disorders of consciousness. Cereb. Cortex *33*, 7193–7210. https://doi.org/10.1093/cercor/bhad031.

34. Dasilva, M., Camassa, A., Navarro-Guzman, A., Pazienti, A., Perez-Mendez, L., Zamora-López, G., Mattia, M., and Sanchez-Vives, M.V. (2021). Modulation of cortical slow oscillations and complexity across anesthesia levels. NeuroImage *224*, 117415. https://doi.org/10.1016/j.neuroimage.2020.117415.

35. Gilson, M., Tagliazucchi, E., and Cofre, R. (2023). Entropy production of Multivariate Ornstein-Uhlenbeck processes correlates with consciousness levels in the human brain. Phys. Rev. E *107*. https://doi.org/10.1103/PhysRevE.107.024121.

36. Ignaccolo, M., Latka, M., Jernajczyk, W., Grigolini, P., and West, B.J. (2010). Dynamics of electroencephalogram entropy and pitfalls of scaling detection. Phys. Rev. E *81*. https://doi.org/10.1103/physreve.81.031909.

37. Gilson, M., Zamora-López, G., Pallarés, V., Adhikari, M.H., Senden, M., Campo, A.T., Mantini, D., Corbetta, M., Deco, G., and Insabato, A. (2020). Model-based whole-brain effective connectivity to study distributed cognition in health and disease. Netw. Neurosci. *4*, 338–373. https://doi.org/10.1162/netn_a_00117.

38. Cugliandolo, L.F., Dean, D.S., and Kurchan, J. (1997). Fluctuation-Dissipation theorems and entropy production in relaxational systems. Phys. Rev. Lett. *79*, 2168–2171. https://doi.org/10.1103/PhysRevLett.79.2168.

39. Cugliandolo, L.F., Kurchan, J., and Parisi, G. (1994). Off equilibrium dynamics and aging in unfrustrated systems. J. Phys. I *4*, 1641–1656. https://doi.org/10.1051/jp1:1994212.

40. Ruiz-Mejias, M., Ciria-Suarez, L., Mattia, M., and Sanchez-Vives, M.V. (2011). Slow and fast rhythms generated in the cerebral cortex of the anesthetized mouse. J. Neurophysiol. *106*, 2910–2921. https://doi.org/10.1152/jn.00440.2011.

41. Sanchez-Vives, M.V. (2020). Origin and dynamics of cortical slow oscillations. Curr. Opin. Physiol. *15*, 217–223. https://doi.org/10.1016/j.cophys.2020.04.005.

42. Camassa, A., Galluzzi, A., Mattia, M., and Sanchez-Vives, M.V. (2022). Deterministic and Stochastic Components of Cortical Down States: Dynamics and Modulation. J. Neurosci. *42*, 9387–9400. https://doi.org/10.1523/JNEUROSCI.0914-22.2022.

43. Garfield, J.M., Garfield, F.B., Stone, J.G., Hopkins, D., and Johns, L.A. (1972). A Comparison of Psychologic Responses to Ketamine and Thiopental–Nitrous Oxide–Halothane Anesthesia. Anesthesiology *36*, 329–338. https://doi.org/10.1097/00000542-197204000-00006.

44. Drummond, J.C., Brebner, J., and Young, P.S. A randomized evaluation of the reversal of ketamine by physostigmine.

45. Massimini, M., Ferrarelli, F., Huber, R., Esser, S.K., Singh, H., and Tononi, G. (2005). Breakdown of Cortical Effective Connectivity during Sleep. Sci. New Ser. *309*, 2228–2232.

46. Sarasso, S., Rosanova, M., Casali, A.G., Casarotto, S., Fecchio, M., Boly, M., Gosseries, O., Tononi, G., Laureys, S., and Massimini, M. (2014). Quantifying Cortical EEG Responses to TMS in (Un)consciousness. Clin. EEG Neurosci. *45*, 40–49. https://doi.org/10.1177/1550059413513723.

47. Koch, C., Massimini, M., Boly, M., and Tononi, G. (2016). Neural correlates of consciousness: progress and problems. Nat. Rev. Neurosci. *17*, 307–321. https://doi.org/10.1038/nrn.2016.22.

48. Engemann, D.A., Raimondo, F., King, J.-R., Rohaut, B., Louppe, G., Faugeras, F., Annen, J., Cassol, H., Gosseries, O., Fernandez-Slezak, D., et al. (2018). Robust EEG-based cross-site and cross-protocol classification of states of consciousness. Brain *141*, 3179–3192. https://doi.org/10.1093/brain/awy251.

49. Schiff, N.D. (2016). Mesocircuit Mechanisms Underlying Recovery of Consciousness Following Severe Brain Injuries: Model and Predictions. In Brain Function and Responsiveness in Disorders of Consciousness, M. M. Monti and W. G. Sannita, eds. (Springer International Publishing), pp. 195–204. https://doi.org/10.1007/978-3-319-21425-2_15.

50. Fischer, C., Luaute, J., and Morlet, D. (2010). Event-related potentials (MMN and novelty P3) in permanent vegetative or minimally conscious states. Clin. Neurophysiol. *121*, 1032–1042. https://doi.org/10.1016/j.clinph.2010.02.005.

51. Höller, Y., Bergmann, J., Kronbichler, M., Crone, J.S., Schmid, E.V., Golaszewski, S., and Ladurner, G. (2011). Preserved oscillatory response but lack of mismatch negativity in patients with disorders of consciousness. Clin. Neurophysiol. *122*, 1744–1754. https://doi.org/10.1016/j.clinph.2011.02.009.

52. Silverstein, B.H., Snodgrass, M., Shevrin, H., and Kushwaha, R. (2015). P3b, consciousness, and complex unconscious processing. Cortex *73*, 216–227. https://doi.org/10.1016/j.cortex.2015.09.004.

53. Forgacs, P.B., Conte, M.M., Fridman, E.A., Voss, H.U., Victor, J.D., and Schiff, N.D. (2014). Preservation of electroencephalographic organization in patients with impaired consciousness and imaging-based evidence of command-following. Ann. Neurol. *76*, 869–879. https://doi.org/10.1002/ana.24283.

54. Muthukumaraswamy, S.D., and Liley, D.Tj. (2018). 1/f electrophysiological spectra in resting and drug-induced states can be explained by the dynamics of multiple oscillatory relaxation processes. NeuroImage *179*, 582–595. https://doi.org/10.1016/j.neuroimage.2018.06.068.

55. Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated information theory: from consciousness to its physical substrate. Nat. Rev. Neurosci. *17*, 450–461. https://doi.org/10.1038/nrn.2016.44.

56. Tononi, G. (2004). An information integration theory of consciousness. BMC Neurosci. *5*, 42. https://doi.org/10.1186/1471-2202-5-42.

57. Tononi, G., Sporns, O., and Edelman, G.M. (1994). A measure for brain complexity: relating functional segregation and integration in the nervous system. Proc. Natl. Acad. Sci. *91*, 5033–5037. https://doi.org/10.1073/pnas.91.11.5033.

58. Andrillon, T., Poulsen, A.T., Hansen, L.K., Léger, D., and Kouider, S. (2016). Neural Markers of Responsiveness to the Environment in Human Sleep. J. Neurosci. *36*, 6583–6596. https://doi.org/10.1523/JNEUROSCI.0902-16.2016.

59. Ferenets, R., Lipping, T., Anier, A., Jantti, V., Melto, S., and Hovilehto, S. (2006). Comparison of Entropy and Complexity Measures for the Assessment of Depth of Sedation. IEEE Trans. Biomed. Eng. *53*, 1067–1077. https://doi.org/10.1109/TBME.2006.873543.

60. Ferenets, R., Vanluchene, A., Lipping, T., Heyse, B., and Struys, M.M.R.F. (2007). Behavior of Entropy/Complexity Measures of the Electroencephalogram during Propofol-induced Sedation: Dose-dependent Effects of Remifentanil. Anesthesiology *106*, 696–706. https://doi.org/10.1097/01.anes.0000264790.07231.2d.

61. Schartner, M., Seth, A., Noirhomme, Q., Boly, M., Bruno, M.-A., Laureys, S., and Barrett, A. (2015). Complexity of Multi-Dimensional Spontaneous EEG Decreases during Propofol Induced General Anaesthesia. PLOS ONE *10*, e0133532. https://doi.org/10.1371/journal.pone.0133532.

62. Maschke, C., O'Byrne, J., Colombo, M.A., Boly, M., Gosseries, O., Laureys, S., Rosanova, M., Jerbi, K., and Blain-Moraes, S. (2024). Critical dynamics in spontaneous EEG predict anesthetic-induced loss of consciousness and perturbational complexity. Commun. Biol. *7*, 946. https://doi.org/10.1038/s42003-024-06613-8.

63. Stikvoort, W., Pérez-Ordoyo, E., Mindlin, I., Escrichs, A., Sitt, J.D., Kringelbach, M.L., Deco, G., and Perl, Y.S. (2024). Nonequilibrium brain dynamics elicited as the origin of perturbative complexity. Preprint at Neuroscience, https://doi.org/10.1101/2024.11.29.625885 https://doi.org/10.1101/2024.11.29.625885.

64. Cloostermans, M.C., Van Meulen, F.B., Eertman, C.J., Hom, H.W., and Van Putten, M.J.A.M. (2012). Continuous electroencephalography monitoring for early prediction of neurological outcome in postanoxic patients after cardiac arrest: A prospective cohort study*. Crit. Care Med. *40*, 2867–2875. https://doi.org/10.1097/CCM.0b013e31825b94f0.

65. Rossi Sebastiano, D., Panzica, F., Visani, E., Rotondi, F., Scaioli, V., Leonardi, M., Sattin, D., D'Incerti, L., Parati, E., Ferini Strambi, L., et al. (2015). Significance of multiple neurophysiological measures in patients with chronic disorders of consciousness. Clin. Neurophysiol. *126*, 558–564. https://doi.org/10.1016/j.clinph.2014.07.004.

66. Rossetti, A.O., Rabinstein, A.A., and Oddo, M. (2016). Neurological prognostication of outcome in patients in coma after cardiac arrest. Lancet Neurol. *15*, 597–609. https://doi.org/10.1016/S1474-4422(16)00015-6.

67. Hofmeijer, J., and Van Putten, M.J.A.M. (2016). EEG in postanoxic coma: Prognostic and diagnostic value. Clin. Neurophysiol. *127*, 2047–2055. https://doi.org/10.1016/j.clinph.2016.02.002.

68. Deco, G., Cruzat, J., Cabral, J., Tagliazucchi, E., Laufs, H., Logothetis, N.K., and Kringelbach, M.L. (2019). Awakening: Predicting external stimulation to force transitions between different brain states. Proc. Natl. Acad. Sci. *116*, 18088–18097. https://doi.org/10.1073/pnas.1905534116.

69. Edlow, B.L., Claassen, J., Schiff, N.D., and Greer, D.M. (2021). Recovery from disorders of consciousness: mechanisms, prognosis and emerging therapies. Nat. Rev. Neurol. *17*, 135–156. https://doi.org/10.1038/s41582-020-00428-x.

70. Sanchez-Vives, M.V., Massimini, M., and Mattia, M. (2017). Shaping the Default Activity Pattern of the Cortical Network. Neuron *94*, 993–1001. https://doi.org/10.1016/j.neuron.2017.05.015.

71. Piarulli, A., Bergamasco, M., Thibaut, A., Cologan, V., Gosseries, O., and Laureys, S. (2016). EEG ultradian rhythmicity differences in disorders of consciousness during wakefulness. J. Neurol. *263*, 1746–1760. https://doi.org/10.1007/s00415-016-8196-y.

72. Mertel, I., Pavlov, Y.G., Barner, C., Müller, F., Diekelmann, S., and Kotchoubey, B. (2020). Sleep in disorders of consciousness: behavioral and polysomnographic recording. BMC Med. *18*, 350. https://doi.org/10.1186/s12916-020-01812-6.

73. Colombo, M.A., Napolitani, M., Boly, M., Gosseries, O., Casarotto, S., Rosanova, M., Brichant, J.-F., Boveroux, P., Rex, S., Laureys, S., et al. (2019). The spectral exponent of the resting EEG indexes the presence of consciousness during unresponsiveness induced by propofol, xenon, and ketamine. NeuroImage *189*, 631–644. https://doi.org/10.1016/j.neuroimage.2019.01.024.

74. Patow, G., Monti, J., Acero-Pousa, I., Idesis, S., Escrichs, A., Perl, Y.S., Ritter, P., Kringelbach, M., Deco, G., and the Alzheimer's Disease Neuroimaging Initiative (2024). Off-Equilibrium Fluctuation-Dissipation Theorem Paves the Way in Alzheimer's Disease Research. Preprint at Neuroscience, https://doi.org/10.1101/2024.09.15.613131 https://doi.org/10.1101/2024.09.15.613131.

75. Tort-Colet, N., Capone, C., Sanchez-Vives, M.V., and Mattia, M. (2021). Attractor competition enriches cortical dynamics during awakening from anesthesia. Cell Rep. *35*, 109270. https://doi.org/10.1016/j.celrep.2021.109270.

76. Pazzini, L., Polese, D., Weinert, J.F., Maiolo, L., Maita, F., Marrani, M., Pecora, A., Sanchez-Vives, M.V., and Fortunato, G. (2018). An ultra-compact integrated system for brain activity recording and stimulation validated over cortical slow oscillations in vivo and in vitro. Sci. Rep. *8*, 16717. https://doi.org/10.1038/s41598-018-34560-y.

77. Massimini, M., and Laureys, S. (2017). Rest EEG recordings in healthy subjects during wakefulness, sleep and anesthesia with ketamine, propofol, and xenon. (Zenodo). https://doi.org/10.5281/ZENODO.806176 https://doi.org/10.5281/ZENODO.806176.

78. Ramsay, M.A.E., Savege, T.M., Simpson, B.R.J., and Goodwin, R. (1974). Controlled Sedation with Alphaxalone-Alphadolone. BMJ *2*, 656–659. https://doi.org/10.1136/bmj.2.5920.656.

79. Murphy, M., Bruno, M.-A., Riedner, B.A., Boveroux, P., Noirhomme, Q., Landsness, E.C., Brichant, J.-F., Phillips, C., Massimini, M., Laureys, S., et al. (2011). Propofol Anesthesia and Sleep: A High-Density EEG Study. Sleep *34*, 283–291. https://doi.org/10.1093/sleep/34.3.283.

80. Rex, S., Schaefer, W., Meyer, P.H., Rossaint, R., Boy, C., Setani, K., Büll, U., and Baumert, J.H. (2006). Positron Emission Tomography Study of Regional Cerebral Metabolism during General Anesthesia with Xenon in Humans. Anesthesiology *105*, 936–943. https://doi.org/10.1097/00000542-200611000-00014.

81. Lee, U., Ku, S., Noh, G., Baek, S., Choi, B., and Mashour, G.A. (2013). Disruption of Frontal–Parietal Communication by Ketamine, Propofol, and Sevoflurane. Anesthesiology *118*, 1264–1275. https://doi.org/10.1097/ALN.0b013e31829103f5.

82. Hockaday, J.M., Potts, F., Epstein, E., Bonazzi, A., and Schwab, R.S. (1965). Electroencephalographic changes in acute cerebral anoxia from cardiac or respiratory arrest. Electroencephalogr. Clin. Neurophysiol. *18*, 575–586. https://doi.org/10.1016/0013-4694(65)90075-1.

83. Estraneo, A., Loreto, V., Guarino, I., Boemia, V., Paone, G., Moretta, P., and Trojano, L. (2016). Standard EEG in diagnostic process of prolonged disorders of consciousness. Clin. Neurophysiol. *127*, 2379–2385. https://doi.org/10.1016/j.clinph.2016.03.021.

84. Reig, R., Mattia, M., Compte, A., Belmonte, C., and Sanchez-Vives, M.V. (2010). Temperature Modulation of Slow and Fast Cortical Rhythms. J. Neurophysiol. *103*, 1253–1261. https://doi.org/10.1152/jn.00890.2009.

85. Sanchez-Vives, M.V., Mattia, M., Compte, A., Perez-Zabalza, M., Winograd, M., Descalzo, V.F., and Reig, R. (2010). Inhibitory Modulation of Cortical Up States. J. Neurophysiol. *104*, 1314–1324. https://doi.org/10.1152/jn.00178.2010.

86. Mattia, M., and Del Giudice, P. (2002). Population dynamics of interacting spiking neurons. Phys. Rev. E *66*, 051917. https://doi.org/10.1103/PhysRevE.66.051917.

87. Comolatti, R., Pigorini, A., Casarotto, S., Fecchio, M., Faria, G., Sarasso, S., Rosanova, M., Gosseries, O., Boly, M., Bodart, O., et al. (2019). A fast and general method to empirically estimate the complexity of brain responses to transcranial and intracranial stimulations. Brain Stimulat. *12*, 1280–1289. https://doi.org/10.1016/j.brs.2019.05.013.

88. Wilson, H.R., and Cowan, J.D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik *13*, 55–80. https://doi.org/10.1007/BF00288786.

89. Wang, M.C., and Uhlenbeck, G.E. (1945). On the Theory of the Brownian Motion II. Rev. Mod. Phys. *17*, 323–342. https://doi.org/10.1103/revmodphys.17.323.

90. Godrèche, C., and Luck, J.-M. (2019). Characterising the nonequilibrium stationary states of Ornstein-Uhlenbeck processes. J. Phys. Math. Theor. *52*, 035002. https://doi.org/10.1088/1751-8121/aaf190.

91. Uhlenbeck, G.E., and Ornstein, L.S. (1930). On the Theory of the Brownian Motion. Phys. Rev. *36*, 823–841. https://doi.org/10.1103/PhysRev.36.823.

92. Hsu, M., and Hsu, D. (2005). A piece-wise harmonic Langevin model of EEG dynamics: Theory and application to EEG seizure detection. Neurocomputing *65–66*, 469–474. https://doi.org/10.1016/j.neucom.2004.11.003.

93. Endemann, C.M., Krause, B.M., Nourski, K.V., Banks, M.I., and Veen, B.V. (2022). Multivariate autoregressive model estimation for high-dimensional intracranial electrophysiological data. NeuroImage *254*, 119057. https://doi.org/10.1016/j.neuroimage.2022.119057.

94. Anderson, C.W., Stolz, E.A., and Shamsunder, S. (1998). Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks. IEEE Trans. Biomed. Eng. *45*, 277–286. https://doi.org/10.1109/10.661153.

95. Gilson, M., Deco, G., Friston, K.J., Hagmann, P., Mantini, D., Betti, V., Romani, G.L., and Corbetta, M. (2018). Effective connectivity inferred from fMRI transition dynamics during

movie viewing points to a balanced reconfiguration of cortical interactions. NeuroImage *180*, 534–546. https://doi.org/10.1016/j.neuroimage.2017.09.061.

96. Deco, G., Sanz Perl, Y., Johnson, S., Bourke, N., Carhart-Harris, R.L., and Kringelbach, M.L. (2024). Different hierarchical reconfigurations in the brain by psilocybin and escitalopram for depression. Nat. Ment. Health *2*, 1096–1110. https://doi.org/10.1038/s44220-024-00298-y.

97. Gilson, M., Moreno-Bote, R., Ponce-Alvarez, A., Ritter, P., and Deco, G. (2016). Estimation of Directed Effective Connectivity from fMRI Functional Connectivity Hints at Asymmetries of Cortical Connectome. PLOS Comput. Biol. *12*, e1004762. https://doi.org/10.1371/journal.pcbi.1004762.

98. Kringelbach, M.L., Perl, Y.S., Tagliazucchi, E., and Deco, G. (2023). Toward naturalistic neuroscience: Mechanisms underlying the flattening of brain hierarchy in movie-watching compared to rest and task. Sci. Adv. *9*, eade6049. https://doi.org/10.1126/sciadv.ade6049.

99. Altaner, B., Polettini, M., and Esposito, M. (2016). Fluctuation-Dissipation Relations Far from Equilibrium. Phys. Rev. Lett. *117*, 180601. https://doi.org/10.1103/PhysRevLett.117.180601.

100. Ohga, N., Ito, S., and Kolchinsky, A. (2023). Thermodynamic Bound on the Asymmetry of Cross-Correlations. Phys. Rev. Lett. *131*, 077101. https://doi.org/10.1103/PhysRevLett.131.077101.