CONSCIOUSNESS 9

A large-scale scientific collaboration has compared two major theories about consciousness. The study, published in *Nature*, involved researchers from around the world for seven years. The theories in question, Integrated Information Theory (IIT) and Global Neuronal Workspace Theory (GNWT), offer distinct visions of how consciousness emerges.

The results suggest that visual areas located at the back of the brain play a crucial role in consciousness. This discovery challenges the idea that the prefrontal cortex would be the main seat of our consciousness. The connections between neurons in visual areas and frontal regions appear essential.

The study involved 256 participants, a record number for this type of research. Scientists used three different methods to measure brain activity. This methodological approach provided robust and comparable data.

The medical implications of these findings are significant. They could help identify signs of consciousness in vegetative state patients. About one quarter of these patients show 'hidden consciousness,' according to a study published in the *New England Journal of Medicine*.

What is Integrated Information Theory (IIT)?

IIT proposes that consciousness emerges when information in the brain is highly integrated and unified. This theory suggests that the degree of consciousness depends on a system's ability to integrate information.

IIT stands out for its mathematical approach to quantifying consciousness. It uses a measure called 'phi' to assess the level of information integration. The higher the phi value, the more conscious the system is considered to be.

This theory has profound implications for understanding altered states of consciousness. It could also apply to artificial systems, raising questions about the possibility of artificial consciousness.

Despite its originality, IIT remains controversial. Some critics point out the difficulty of objectively measuring consciousness and applying the theory to non-biological systems.

How does Global Neuronal Workspace Theory (GNWT) work?

GNWT postulates that consciousness results from the broadcasting of information across a network of brain regions. According to this theory, certain information is 'brought into the light' and becomes accessible to various parts of the brain.

This broadcasting allows flexible integration of information, essential for complex tasks. GNWT emphasizes the role of the prefrontal cortex in this process, considered the conductor of consciousness.

The theory explains how information can transition from unconscious to conscious. It provides a framework for understanding phenomena like attention and working memory.

However, GNWT still needs to explain exactly how information is selected for broadcasting. The precise mechanisms of this selection remain an active research topic.

- Published: 30 April 2025

Adversarial testing of global neuronal workspace and integrated information theories of consciousness

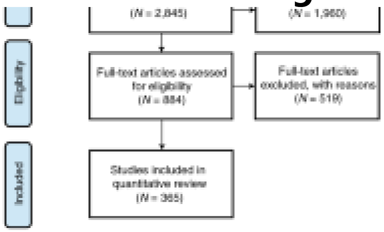
- Cogitate Consortium,
- Oscar Ferrante,
- Urszula Gorska-Klimowska,
- Simon Henin,
- Rony Hirschhorn,
- Aya Khalaf,
- Alex Lepauvre,
- Ling Liu,
- David Richter,
- Yamil Vidal,
- Niccolò Bonacchi,
- Tanya Brown,
- Praveen Sripad,
- Marcelo Armendariz,
- Katarina Bendtz,
- Tara Ghafari,
- Dorottya Hetenyi,
- Jay Jeschke,
- Csaba Kozma,
- David R. Mazumder,
- Stephanie Montenegro,
- Alia Seedat,
- Abdelrahman Sharafeldin,
- Shujun Yang,
- Lucia Melloni

Nature (2025)**56k** Accesses

Abstract

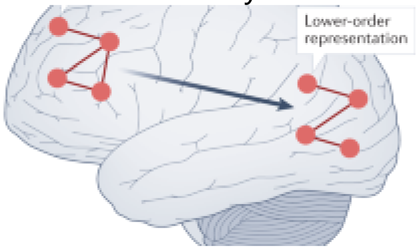
Different theories explain how subjective experience arises from brain activity[1,2]. These theories have independently accrued evidence, but have not been directly compared[3]. Here we present an open science adversarial collaboration directly juxtaposing integrated information theory (IIT)[4,5] and global neuronal workspace theory (GNWT)[6,7,8,9,10] via a theory-neutral consortium[11,12,13]. The theory proponents and the consortium developed and preregistered the experimental design, divergent predictions, expected outcomes and interpretation thereof[12]. Human participants ($n = 256$) viewed suprathreshold stimuli for variable durations while neural activity was measured with functional magnetic resonance imaging, magnetoencephalography and intracranial electroencephalography. We found information about conscious content in visual, ventrotemporal and inferior frontal cortex, with sustained responses in occipital and lateral temporal cortex reflecting stimulus duration, and content-specific synchronization between frontal and early visual areas. These results align with some predictions of IIT and GNWT, while substantially challenging key tenets of both theories. For IIT, a lack of sustained synchronization within the posterior cortex contradicts the claim that network connectivity specifies consciousness. GNWT is challenged by the general lack of ignition at stimulus offset and limited representation of certain conscious dimensions in the prefrontal cortex. These challenges extend to other theories of consciousness that share some of the predictions tested here[14,15,16,17]. Beyond challenging the theories, we present an alternative approach to advance cognitive neuroscience through principled, theory-driven, collaborative research and highlight the need for a quantitative framework for systematic theory testing and building.

Similar content being viewed by others



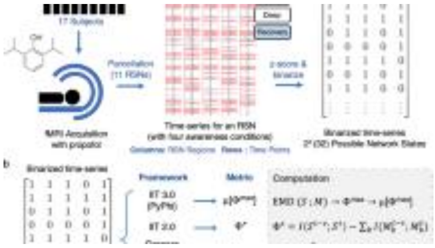
The ConTraSt database for analysing and comparing empirical studies of consciousness theories

Article 21 February 2022



Theories of consciousness

Article 03 May 2022



An implementation of integrated information theory in resting-state fMRI

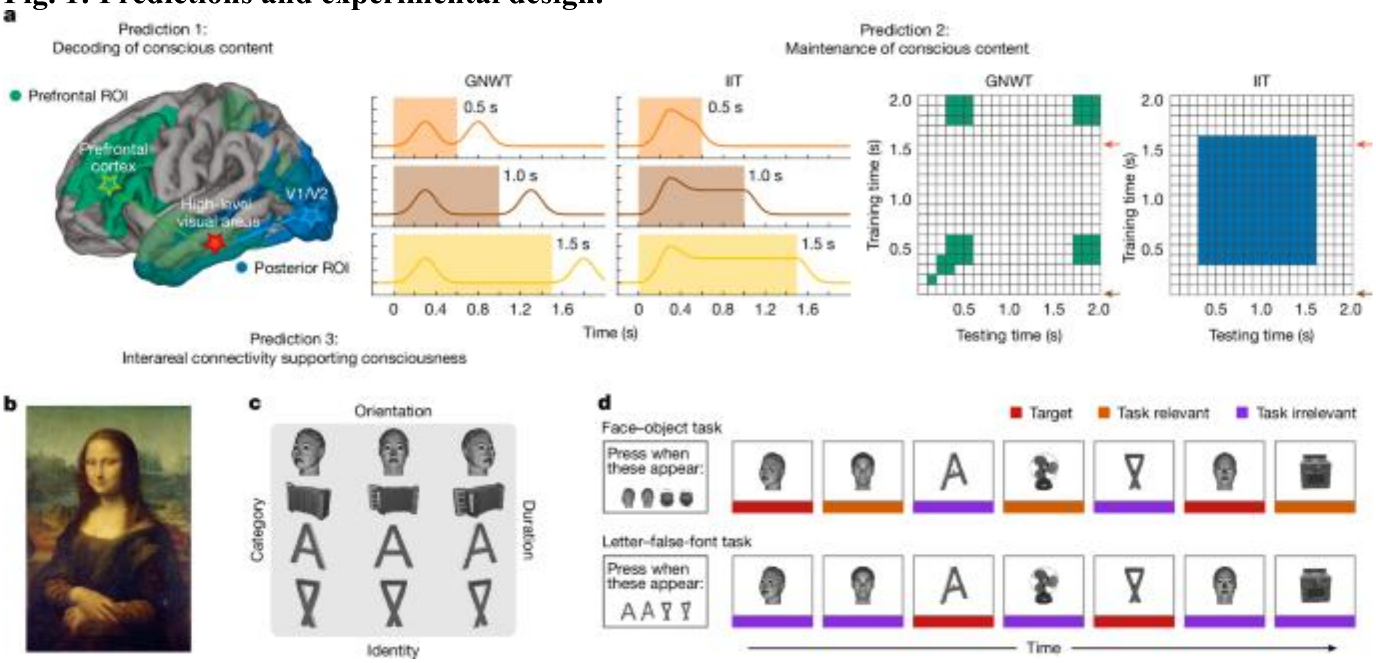
Article Open access 05 July 2023

Main

Philosophers and scientists have sought to explain the subjective nature of consciousness (for example, the feeling of pain or of seeing a colourful rainbow) and how it relates to physical processes in the brain[18,19]. This quest has led to various theories of consciousness evolving in parallel[1,2,3] and often providing incompatible accounts of the neural basis of consciousness[1,2]. Furthermore, empirical support for a given theory is often highly dependent on methodological choices, pointing towards a confirmation bias in theory testing[3]. Convergence on a broadly accepted neuroscientific theory of consciousness will have profound medical, societal and ethical implications.

To advance this goal, we tested two theories of consciousness, through a large-scale, open-science adversarial collaboration[11,12,20,21,22] aimed at accelerating progress in consciousness research by building on constructive disagreement. We brought together proponents of IIT[4,5] and GNWT[6,23], in addition to theory-neutral researchers. The group identified differential existing and novel predictions of the two theories and developed an experimental design to test them (Fig. 1a). We preregistered these predictions, including pass or fail criteria, expected outcomes and their interpretation ex-ante[11,12]. We focus on IIT and GNWT, among other widely discussed theories (for example, recurrent processing theory and higher-order theories[1,2]), because they feature prominently in consciousness science, as demonstrated by a recent systematic review of the literature[3].

Fig. 1: Predictions and experimental design.



a, Predictions of IIT and GNWT. For prediction 1 (decoding of conscious content), IIT predicts maximal decoding of conscious content in posterior brain areas, whereas GNWT emphasizes a necessary role for the PFC. For prediction 2 (maintenance of conscious content), IIT posits that conscious content is actively maintained in the posterior cortex, whereas GNWT predicts brief content-specific ignition (approximately 0.3–0.5 s) in the PFC at stimulus onset and offset, with content stored in a non-conscious silent state between these events. Waveforms (left) and temporal generalization matrices (right) illustrate the predicted amplitude-based and information-based temporal profiles: coloured rectangles indicate the three stimulus durations for PFC (GNWT) and posterior cortex (IIT; left); the arrows mark stimulus onset (brown) and offset (red), whereas predicted temporal generalization is depicted in green (GNWT) and blue (IIT; right). For prediction 3 (interareal connectivity supporting consciousness), the stars and arrows on the brain diagram illustrate predicted synchrony patterns, with green representing GNWT and blue representing IIT. Brain surface is from Freesurfer. **b**, Conscious experience is multifaceted. For instance, viewing the Mona Lisa involves experiencing it as occupying a specific spatial location, categorizing it as a face, recognizing an identity and noting its leftward orientation, with this complex experience maintained over time. **c**, To manipulate conscious content, stimuli varied across four dimensions: category (faces, objects, letters and false fonts), identity (different exemplars within each category), orientation (left, right and front views) and duration (0.5 s, 1.0 s and 1.5 s). Example stimuli are shown. **d**, Experimental paradigm. Participants detected predefined targets (for example, a specific face and object or a letter and false font) in sequences of single, high-contrast stimuli. Each trial contained three stimulus types: targets (red; coloured frames for illustration only), task-relevant stimuli (orange-red; same categories as targets) and task-irrelevant stimuli (purple; other categories). Blank intervals between stimuli are not depicted. Object stimulus images in panels **c**,**d** are courtesy of Michael J. Tarr, Carnegie Mellon University, http://www.tarrlab.org/; face stimuli were created using FaceGen Modeler 3.1.

IIT and GNWT explain consciousness differently: IIT proposes that consciousness is the intrinsic ability of a neuronal network to influence itself, as determined by the amount of maximally irreducible integrated information (phi) supported by a network in a state. On the basis of theoretical and neuroanatomical considerations, IIT suggests that a complex of maximum phi probably resides primarily in the posterior cerebral cortex, in a temporo–parietal–occipital 'hot zone'[4,5,24,25]. GNWT instead posits that consciousness arises from global broadcasting and late amplification (or 'ignition') of information across interconnected networks of higher-order sensory, parietal and especially prefrontal cortex (PFC)[6,9,23].

Both theories have a mathematical or computational core (integrated information for IIT and global workspace for GNWT) and proposed biological implementations (posterior cortex versus PFC and associated areas, respectively). Although it is difficult to test the mathematical or computational core of either theory directly, their competing biological implementations are empirically testable with current methodologies. Thus, our study focuses on brain regions where the predictions diverge most notably—posterior cortex for IIT and PFC for GNWT, rather than the associated areas in higher-order sensory or parietal cortex—to facilitate maximally diagnostic experiments.

One consequence of this biological focus is that theorists could respond to challenging data by modifying the proposed biological implementation while retaining the mathematical or computational core of a theory. Another consequence is that some predictions (and their associated consequences) may overlap with other theories of consciousness that share similar biological bases, such as higher-order theories[16,17] in the PFC and local recurrency theories[14,26] in the visual cortex. Although these are inherent aspects of studying theoretical proposals about neural mechanisms of consciousness, the results are expected to help the community make more informed judgements about the tested theories (for rationale, see the preregistration document[27]).

Preregistered predictions and analyses

We tested three preregistered, peer-reviewed predictions of IIT and GNWT[12] for how the brain enables conscious experience (Fig. 1a). Prediction 1 addresses the cortical areas holding information about different aspects of conscious content. IIT predicts that conscious content is maximal in posterior brain areas, whereas GNWT predicts a necessary role for PFC. Prediction 2 pertains to the maintenance of conscious percepts over time[28,29,30]. IIT predicts that conscious content is actively maintained by neural activity in the posterior 'hot zone' throughout the duration of a conscious experience, whereas GNWT predicts ignition events in PFC at stimulus onset and offset, updating the global workspace, with activity-silent information maintenance in between[31]. Prediction 3 examines interareal connectivity during conscious perception. IIT predicts sustained short-range connectivity within the posterior cortex, linking low-level sensory (V1/V2) with high-level category-selective areas (for example, fusiform face area and lateral occipital cortex), whereas GNWT predicts long-range connectivity between high-level category-selective areas and PFC. The combination of predictions, tested through highly powered, multimodal studies, places a high bar for either theory to pass, rendering failures more informative. Predictions were differentially weighted on the basis of their centrality to the theory and methodological considerations (Extended Data Table 1; for an additional preregistered non-critical analysis, see section 8 in Supplementary Information).

To empirically test these predictions, we investigated the content and temporal extent of conscious visual experiences, focusing on their phenomenological richness and multifaceted nature, even for a single stimulus. For instance, when viewing the Mona Lisa (Fig. 1b), one experiences it as having a specific identity, orientation and location in visual space for as long as one looks at the painting. To approximate such multifaceted experiences, we manipulated several attributes of conscious content by presenting suprathreshold visual stimuli across four different categories (faces, objects, letters and false fonts), each containing 20 unique identities shown in three orientations (front, left and right view) and for three durations (0.5, 1.0 and 1.5 s). In each block, participants were instructed to detect two infrequent target stimuli from either the pictorial (face–object) or symbolic (letter–false fonts) stimulus categories (for example, a specific face or object), making these categories task relevant for that block (Fig. 1c,d).

This paradigm offers several advantages. First, it provides robust conditions to test the predictions of the theories by focusing on clearly experienced conscious content, studied through a high signal-to-noise, suprathreshold, fully attended single stimulus at fixation. This amplifies the significance of any challenges to the theories, as they cannot be explained by weak signals. Second, it minimizes task and report confounds, isolating neural activity specifically related to consciousness. Third, it allows testing of novel predictions to address previously unexplored questions, that is, how experience is maintained over time, refining theories and yielding new insights.

All research was conducted by theory-neutral teams to minimize confirmatory bias. We evaluated the predictions of theories in 256 participants performing the same behavioural task in three neuroimaging modalities: functional magnetic resonance imaging (fMRI; $n = 120$), magnetoencephalography (MEG; $n = 102$) and intracranial electroencephalography (iEEG; $n = 34$). To overcome the spatial and temporal limitations of each modality, we combined whole-brain, non-invasive fMRI and MEG with invasive iEEG, ensuring methodological rigour. Combined with large sample sizes, this minimizes the likelihood that negative results are due to methodological or sensitivity issues. Data collection occurred in two (or three) independent laboratories for each modality to guarantee generalization across groups of participants, instruments and experimenters. To foster informativeness, reproducibility and robustness, we (1) separated theory proponents from data acquisition and analysis to minimize bias and post-hoc interpretation, (2) used a multimodal approach that maximizes spatiotemporal resolution and coverage for a stringent and comprehensive tests of the theories in humans, (3) predefined large samples to increase statistical power, (4) followed standardized[32] and preregistered protocols[12] to reduce setup differences and confirmatory bias[22] (see sections 1 and 2 in Supplementary Information), and (5) implemented an analysis optimization phase (one-third of the sample) followed by a final testing phase (two-thirds of the sample) on independent data for result validation[33]. Consequently, this large-scale international effort aimed at implementing a rigorous adversarial collaboration framework, thereby establishing a precedent for an alternative scientific approach.

Decoding of conscious content

According to IIT, PFC is not necessary for consciousness. Consequently, decoding conscious content should be most effective from the posterior cortex, and adding PFC activity as additional information should not improve decoding accuracy. This prediction was considered non-critical for testing IIT, as the theory focuses on the intrinsic, causal perspective of information within a neural substrate rather than the amount of information decodable from the perspective of an extrinsic observer[5]. By contrast, GNWT posits that conscious content can be decoded from PFC activity. Both theories predict that conscious content should be evident in theory-relevant areas independently of other cognitive processes (for example, report and task); thus, conscious content should be present irrespective of task manipulations[34,35]. This prediction was tested by evaluating the decoding accuracy of stimulus category (faces–objects (pictorial) and letters–false fonts (symbolic)) and orientation (left, right and front facing) in all theory-relevant areas. All stimulus categories alternated between being task-relevant and task-irrelevant across blocks (Fig. 1d). Stimulus orientation, being orthogonal to the task, remained task-irrelevant in all blocks.

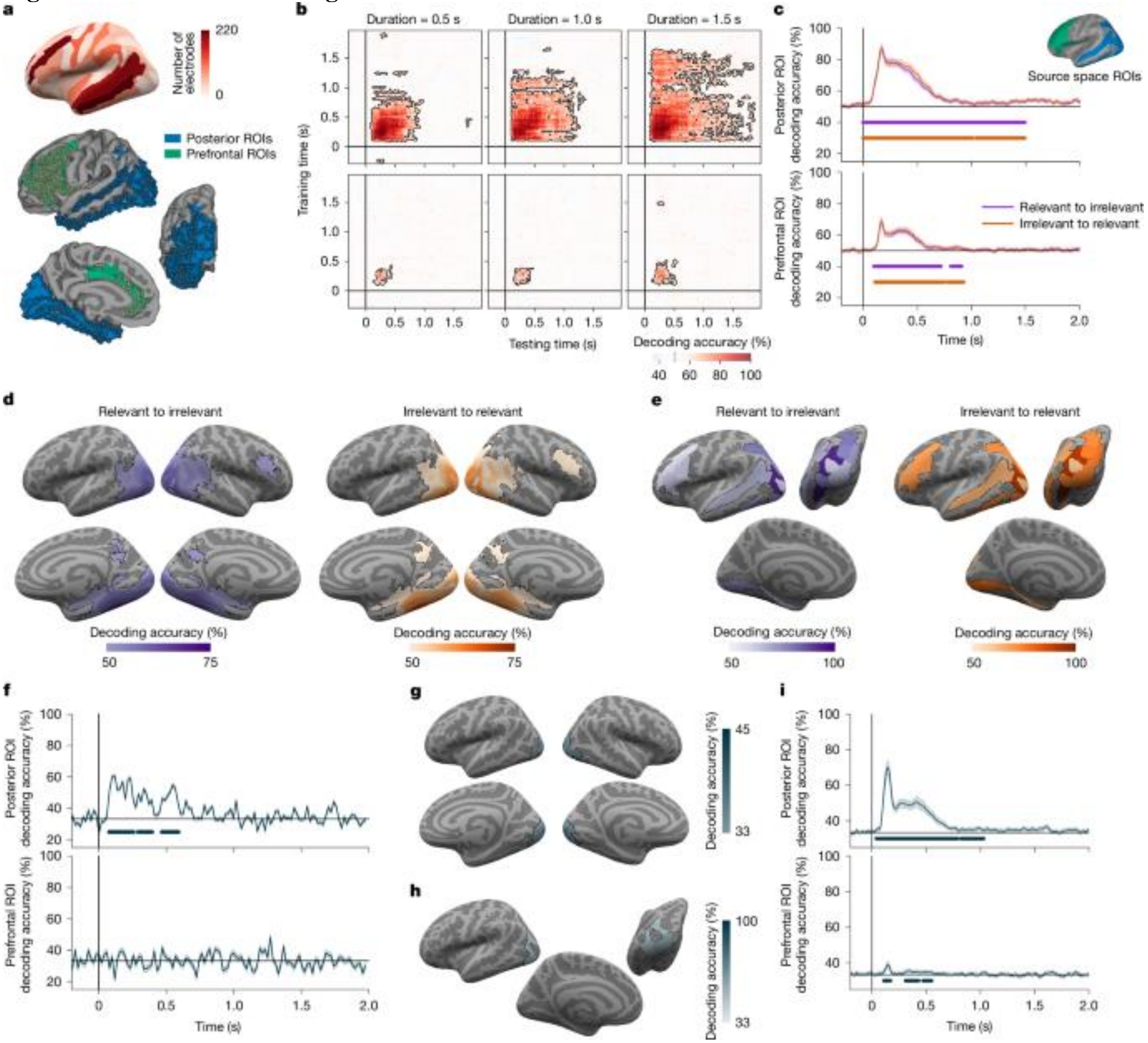
On the basis of our preregistered predictions and pre-approved interpretations[27] (Extended Data Table 1), the theories would pass the test if decoding is possible for both category (in at least one category pairing) and orientation (in at least one category), but would fail otherwise. Testing both category and orientation decoding constitutes a stringent test, as it requires two conditions to be satisfied, increasing the likelihood of failures[36], while capturing the critical multidimensionality of conscious content, that is, phenomenological richness (Fig. 1b).

For decoding of category, we tested whether information is present in the relevant regions irrespective of the task, using a cross-task generalization approach (see Methods).

Here we report the most robust results for decoding of category (faces–objects) and orientation (left, right and front views of faces). Qualitatively similar results were observed for decoding of letters–false fonts (Extended Data Fig. 1a–d). Results for orientation decoding were consistent across stimulus categories and data modalities in posterior cortex, but mostly absent in PFC (see section 5.1.2 in Supplementary Information).

In the iEEG data, pattern classifiers were trained on high gamma frequency band (70–150 Hz), which correlates with spiking activity[37,38], at each time point in the task-irrelevant condition, and tested in the task-relevant condition, for each stimulus duration and category, and across all electrodes within the theory-relevant region of interests (ROIs; see Fig. 2a for a visualization of ROIs and Methods for anatomical ROI definitions). In the posterior ROIs, face–object decoding showed significant cross-task generalization (more than 95% accuracy) for the approximate duration of the stimulus (Fig. 2b, top row). In the PFC ROIs, significant cross-task face–object decoding accuracy (approximately 70%) was also evident, but the temporal generalization of this decoding was restricted to approximately 0.2–0.4 s (Fig. 2b, bottom row). Training on task-relevant and testing on task-irrelevant trials showed similar results (Extended Data Fig. 1e; within-task decoding is presented in Extended Data Fig. 2).

Fig. 2: Prediction 1: decoding of conscious content.



a, Spatial coverage of intracranial electrodes ($n_{\text{patients}} = 29$) on a standard inflated cortical surface map (top), and within theory-defined ROIs (bottom): posterior (blue; $n_{\text{electrodes}} = 583$) and prefrontal (green; $n_{\text{electrodes}} = 576$). **b**, iEEG cross-task temporal generalization of decoding of high-gamma signal. Pattern classifiers were trained to discriminate stimulus category (faces–objects) in the task-irrelevant condition at each time point and tested in the task-relevant condition across all time-points. Columns denote stimulus durations (0.5 s (left), 1.0 s (centre) and 1.5 s (right)), and rows indicate theory ROIs (posterior (top) and prefrontal (bottom)). Contoured red-shaded regions depict significant above-chance (50%) decoding. Here and below, significance was evaluated through a non-parametric cluster-based permutation test ($P < 0.05$; two-sided). **c**, MEG average cross-task decoding of stimulus category ($n = 65$) from task-relevant to task-irrelevant stimuli (purple) and vice versa (orange), separately for the posterior (top) and prefrontal (bottom) ROIs, depicted on inflated cortical surfaces (posterior in blue and prefrontal in green), across durations, using pseudotrial aggregation. Underlying lines indicate significance. The shading depicts 95% CI across participants. **d**, fMRI searchlight cross-task decoding of stimulus category ($n = 73$), collapsed across durations, from task-relevant stimuli to task-irrelevant stimuli (left; purple) or vice versa (right; orange). The outlined coloured regions on the inflated cortical surfaces (left–right lateral views; right–left medial views (bottom)) indicate significant above-chance decoding. **e**, iEEG ROIs significant cross-task decoding of stimulus category, collapsed across durations. Conventions are as in panel **d**, displayed from a left lateral (top left), posterior (top right) and left medial (bottom) views. **f**, iEEG average decoding of stimulus orientation (left, right and front) within posterior (top) and prefrontal (bottom) ROIs, collapsed across durations. Underlying lines indicate above-chance (33%) decoding. The shading depicts 95% CI estimated across cross-validation folds. **g**, fMRI searchlight decoding of face orientation (left, right and front). Regions with significantly above-chance (33%) decoding accuracies are outlined in blue. **h**, iEEG ROIs decoding of face orientation (left, right and front). Conventions are as in panel **g**. **i**, MEG ROIs average decoding of face orientation (left, right and front). Conventions are as in panel **f**. Brain surfaces in panels **a**,**c**–**e**,**g**,**h** are from Freesurfer.

Although electrode coverage across our sample of iEEG patients ($n = 29$ for decoding analyses) was exceptional in the relevant brain regions (Fig. 2a; PFC ROIs $n_{\text{electrodes}} = 576$, posterior ROIs $n_{\text{electrodes}} = 583$), we further analysed a larger population of healthy participants ($n = 65$) using MEG, focusing on theory-relevant ROIs (see Methods). Here too, cross-task generalization of face–object decoding was significant in both posterior and prefrontal ROIs (Fig. 2c) within the theory-predicted time windows. Cross-temporal generalization of decoding in MEG was sustained in posterior ROIs and brief in PFC ROIs for all three stimulus durations (see section 5.1.1.2 in Supplementary Information).

We leveraged the higher spatial resolution of fMRI ($n = 73$) to complement the analysis. A searchlight approach (see Methods) revealed distributed, robust cross-task generalization (approximately 75%) in the striate and extrastriate, ventral temporal and intraparietal cortex (Fig. 2d and Extended Data Table 2). Generalization in PFC had lower accuracy (approximately 60%) and was spatially restricted to middle and inferior frontal cortex regions (Fig. 2d). Theory-relevant ROIs defined in the Destrieux atlas yielded comparable results (see section 5.1.1.3 in Supplementary Information). These results closely matched those from iEEG-restricted to theory-specified ROIs and time windows (Fig. 2e). Hence, across modalities, face–object decoding occurred in both posterior and prefrontal ROIs, consistent with IIT and GNWT predictions.

Given the rich and multidimensional nature of conscious content, we assessed the decoding of stimulus orientation, which was always task-irrelevant. We obtained divergent results for IIT and GNWT: decoding of face orientation (left, right or front views) was achieved in posterior but not in prefrontal ROIs, both with iEEG (Fig. 2f,h, approximately 95% with pseudotrial aggregation; Extended Data Fig. 3a) and the fMRI searchlight approach (Fig. 2g, approximately 45%). Decoding of face orientation was robust from MEG cortical time series in posterior ROIs (approximately 75% with pseudotrial aggregation), but was weaker, yet above chance (35%), in prefrontal ROIs (Fig. 2i), with a possibility of signal leakage from posterior regions (Extended Data Fig. 3b). Bayesian testing further validated these findings. For iEEG, Bayes factor $(\text{BF})_{01}$ values (5.11–8.65) supported the null hypothesis of no face orientation decoding in prefrontal regions. fMRI Bayesian analysis revealed substantial-to-very-strong support for the null hypothesis in 34–55% of prefrontal voxels (BF_{01}: 3–71.5), with support for the alternative hypothesis in only 1–9% of voxels, whereas the rest remained inconclusive. Across all modalities, orientation decoding was observed for letters and false fonts—but not objects—in posterior, but not prefrontal, ROIs (see section 5.1.2 in Supplementary Information).

Finally, we tested IIT's prediction that prefrontal regions do not contribute further information beyond that specified by posterior areas (or may even degrade performance by introducing noise)[39]. If PFC activity increased decoding accuracy, IIT would be challenged, whereas no improvement in decoding accuracy would align with both IIT and GNWT, as GNWT posits that PFC workspace neurons broadcast but do not add information. We compared the performance of decoders trained exclusively on posterior ROIs with those trained on posterior and prefrontal ROIs together (Extended Data Fig. 3c; see Methods). Across critical time-resolved methods (iEEG and MEG) and various PFC ROI definitions, adding prefrontal ROIs did not improve—and in some cases reduced—category and orientation decoding (Extended Data Fig. 3d,e and see section 5.1.3 in Supplementary Information for non-critical fMRI results). Bayesian testing confirmed these findings: we found strong evidence against increased decoding accuracy when including PFC ROIs for category decoding (face–object: iEEG $BF_{01} = 1.94 \times 10^4$ and MEG $BF_{01} = 3.05$; letter–false font: iEEG $BF_{01} = 1.91 \times 10^5$ and MEG $BF_{01} = 4.70$) and face orientation (iEEG $BF_{01} = 1{,}205$ and MEG $BF_{01} = 3.26$).

Maintenance of conscious content

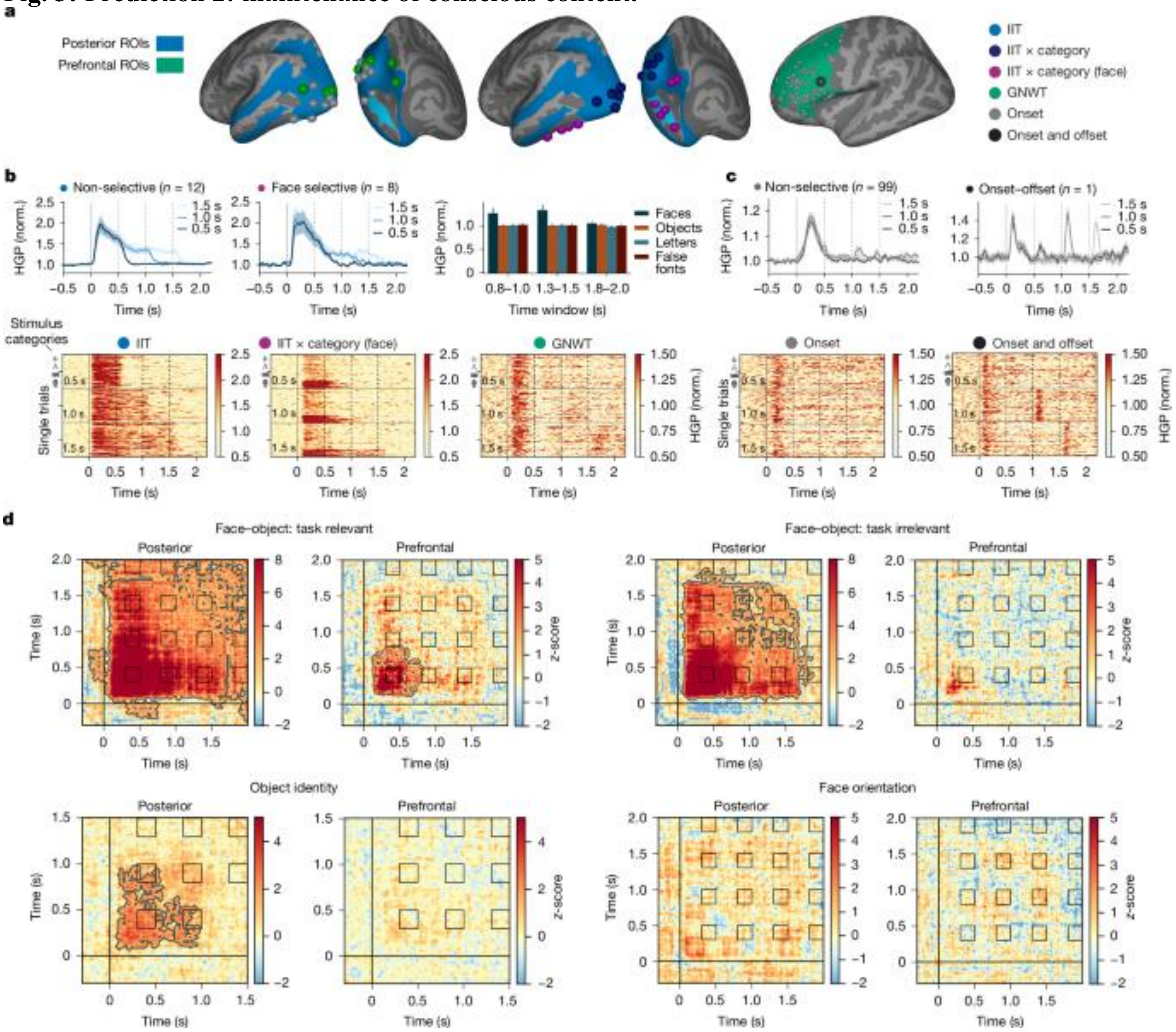
According to IIT, the network that specifies the content of consciousness in posterior cortex is actively maintained over the duration of the conscious experience (manipulated here via different stimulus durations). By contrast, GNWT predicts brief, content-specific ignition in PFC within 0.3–0.5 s after stimulus onset, as the workspace is updated[12]. Activity then decays back to baseline, with information maintained in a latent state, until another ignition marks the offset of the current percept and the onset of a new percept (for example, the fixation screen following stimulus offset). Thus, although the underlying brain responses (the workspace update) are temporally discrete (that is, an onset and an offset response), the conscious experience can be temporally continuous, spanning from one update to the next.

Following the preregistration[27] (Extended Data Table 1), IIT would be challenged if sustained content-specific information and activation tracking stimulus duration were absent in the posterior cortex. GNWT would be challenged if transient prefrontal activation (at stimulus onset and offset) was not observed. These patterns were expected for at least one conscious feature (category, identity or orientation). We assessed activation strength as a function of stimulus duration and the informational content of this activation within theory-relevant ROIs. For IIT, both activation and information content were critical predictions, jointly determining result interpretation. For GNWT, activation was the primary measure owing to the difficulty of reliably detecting content-specific reinstatement at stimulus offset. The temporal predictions were tested in time-resolved iEEG and MEG data. We focused on the task-irrelevant condition as it best isolates neural activity related to consciousness while minimizing confounds (see sections 6.1 and 6.2.9 in Supplementary Information for task-relevant results).

First, we tested the predictions of the theories by investigating iEEG neural activation as a function of stimulus duration using linear mixed models (LMMs; see Methods) to model the time course of neural activity in the high gamma. Among the 31 patients included, 29.5% (194 of 657) of posterior ROI electrodes and 18.7% (123 of 655) of PFC ROI electrodes exhibited high gamma responses to stimuli (see section 6.1.2 in Supplementary Information).

In posterior cortex ROIs, 25 electrodes (out of 657) measured sustained activity tracking stimulus duration (Extended Data Table 3 for electrode localization and section 6.1.1 in Supplementary Information for results of the full model), consistent with IIT's (Fig. 3a). Of these, 12 electrodes tracked duration independent of stimulus category, primarily in early visual areas (for example, occipital pole; Fig. 3b), whereas 13 showed category-specific duration tracking (mostly for faces) in the ventral temporal cortex (for example, lateral fusiform gyrus; Fig. 3b). Overall, only a small proportion of electrodes exhibited both category selectivity and duration tracking—for example, just 15% (8 of 53) of face-selective electrodes showed sustained activity as predicted by IIT, suggesting a sparse neural substrate. These responses were mostly localized to the lateral fusiform gyrus, whereas the majority of face-selective electrodes displayed transient activations at stimulus onset across the striate, extrastriate and ventral areas (see section 6.1.2 in Supplementary Information).

Fig. 3: Prediction 2: maintenance of conscious content.



a, Intracranial electrode localization on the MNI template, for posterior (left; blue; $n_{\text{patients}} = 31$ and $n_{\text{electrodes}} = 657$) and prefrontal (right; green; $n_{\text{patients}} = 31$ and $n_{\text{electrodes}} = 655$) ROIs. Electrodes are colour coded by response type based on model comparison (see Methods): sustained non-category-selective activation (light blue; $n = 12$), sustained category-selective activation (dark blue; $n = 5$), sustained face-selective activation (purple; $n = 8$), biphasic onset–offset activation in posterior areas (green; $n = 11$) and in PFC (black; $n = 1$), and onset-responsive activation in PFC (grey; $n = 99$). Brain surfaces are from Freesurfer. **b**, Posterior ROI activation. Time-series plots depict average high gamma (HG), separated by stimulus duration (0.5 s (dark), 1 s (medium) and 1.5 s (light)) for non-selective (left) and face-selective (middle) electrodes. The shading here and in panel **c** depicts standard error of the mean across electrodes and trials. The barplots (right) depict the average HG signal across sustained face-selective electrodes ($n = 8$) in 1.5-s trials, separated by category (faces in dark blue, objects in orange, letters in turquoise and false fonts in dark red) and theory-defined time windows (*x* axis). Raster plots show single-trial ($n = 320$) HG of individual electrodes during task-irrelevant trials: a sustained non-selective (left), sustained face-selective (middle) and onset–offset (right) electrode. The rows depict single trials, sorted per duration (from top: 0.5, 1.0 and 1.5 s), and then category (from top: false fonts, letters, objects and faces). **c**, Prefrontal ROI activation. Time-series plots (top left) depict the average HG response per stimulus duration (shades of grey) for onset-responsive electrodes ($n = 99$) in task-irrelevant trials ($n = 320$). Average HG response per stimulus duration for a single electrode exhibiting onset–offset responses, with an earlier-than-predicted offset (top right). Raster plots for example onset (bottom left) and onset–offset (bottom right) responses are also shown. Conventions are as in panel **b**. **d**, Cross-temporal RSA matrices in posterior ($n_{\text{patients}} = 28$ and $n_{\text{electrodes}} = 583$) and prefrontal ($n_{\text{patients}} = 28$ and $n_{\text{electrodes}} = 576$) ROIs. Titles indicate the compared contrasts, and subtitles denote the ROIs. Matrix values represent z-scored, within-class-corrected correlation distances derived from a label shuffle null distribution. Contours denote significance (cluster-based permutation tests, $P < 0.05$, upper tail).

In PFC ROIs, 99 and 24 electrodes showed non-selective and category-selective onset responses, respectively (Fig. 3c). However, none of the 655 electrodes measured the temporal profile predicted by GNWT (that is, onset and offset). Bayesian analysis confirmed this result ($BF_{01} > 3$ for all electrodes in PFC ROIs), providing stronger evidence for either an intercept-only or a time-varying amplitude model over the GNWT model, with or without category interaction. This null result was not due to analysis limitations, as the LMM successfully detected the GNWT-predicted pattern in 10 electrodes in other ROIs (in the striate or extrastriate cortex; Fig. 3b). An exploratory decoding analysis of stimulus duration with unrestricted temporal profiles identified only one electrode, in the inferior frontal sulcus, showing the GNWT-predicted pattern, although with transient responses occurring earlier than expected (0.15 s post-onset and post-offset; Fig. 3c). Additional control analyses confirmed the IIT-predicted pattern in posterior ROIs and the absence of the GNWT-predicted pattern in PFC ROIs (see sections 6.2.1–6.2.3 in Supplementary Information).

We used LMMs to track gamma frequency band (60–90 Hz) power changes from the MEG source time series across posterior (15 parcels) and PFC (11 parcels) ROIs. Although signals were strong in posterior areas, none of the theory-based models adequately fit the data ($BF_{01} > 3$ for all parcels; see section 6.1.3.1 in Supplementary Information). We also examined alpha band activity (8–13 Hz), which negatively correlates with neural spiking activity[40,41]. Validation of theoretical predictions from iEEG and MEG data was inconclusive: no prefrontal iEEG electrodes showed the GNWT-predicted combination of an onset and offset response ($BF_{01} > 3$ for all prefrontal electrodes); instead, this pattern appeared in posterior sites and in MEG data, including the anterior cingulate cortex. However, the MEG results were sensitive to parameter choices, and signal leakage from posterior sites could not be ruled out (see sections 6.1.1 and 6.1.3.2 in Supplementary Information).

Next, we used cross-temporal representational similarity analysis (RSA) on both iEEG and MEG source data, within each theory-relevant ROI, to test the predictions of IIT and GNWT about the temporal profile of the maintenance of conscious content (Fig. 1a, middle panel): sustained versus phasic (onset and offset) representation for IIT and GNWT, respectively. This test was critical for IIT only. Results for faces and objects are presented below (see Extended Data Fig. 4 for similar results for letters–false fonts).

In iEEG, we calculated the correlation distance between high gamma activity patterns across 583 electrodes in posterior ($n_{\text{patients}} = 28$) and 576 electrodes in PFC ($n_{\text{patients}} = 28$) ROIs separately (see Methods). We analysed the 1.5-s duration trials because this condition provided the strongest contrast between the temporal profiles predicted by the theories.

In posterior cortex ROIs, cross-temporal RSA revealed sustained face–object categorical representation, with larger correlation distances between categories (face–objects) than within category (face, object; compare Fig. 3d, left, with Fig. 1a). The RSA matrix matched the IIT model better than the GNWT model (see section 6.3 in Supplementary Information for results of all contrasts).

In PFC ROIs, cross-temporal RSA revealed transient face–object categorical representation at stimulus onset, but not at stimulus offset. Consequently, no significant correlation was found with the GNWT onset and offset model (compare Fig. 3d, right, with the predicted pattern in Fig. 1a). This pattern held even for the task-relevant condition, in which face–object information was stronger, more stable and longer lasting. Additional evidence for the absence of GNWT-predicted patterns in PFC ROIs emerged from three control analyses using (1) feature selection, which improved RSA in PFC; (2) modified time-windows to account for a potential earlier ignition at stimulus offset; and (3) a decoding analysis time-locked to stimulus offset to enhance sensitivity (see section 6.4 in Supplementary Information). These results align with two independent studies using comparable methods[42,43].

It has been argued that because conscious experiences are specific, the representation of identity and orientation are more stringent tests of the neural substrate of conscious experience[44] than of category. We thus also evaluated the predictions of the theories on these dimensions.

In posterior ROIs, iEEG revealed sustained object identity information, with smaller distances for same-identity objects than for different identities (Fig. 3d). The IIT model significantly correlated with the observed RSA matrix, providing a better fit than the GNWT model. Similar results were found for letter and false-font identity, but not for faces (Extended Data Fig. 4). In PFC ROIs, identity information was absent for all categories across analysed time windows (Fig. 3d, objects). Face orientation information appeared weakly in posterior ROIs at stimulus onset but was not sustained, decaying after 0.5 s, contrary to the predictions of IIT. No face orientation information was detected in PFC ROIs (Fig. 3d). Finally, the predictions of neither theory were supported for category, identity or orientation by the MEG data (see section 6.5 for Supplementary Information).

Interareal connectivity

IIT predicts sustained gamma-band connectivity within the posterior cortex, that is, between high-level and low-level sensory areas (V1/V2), throughout any conscious visual experience. By contrast, GNWT predicts brief, late-phase metastable connectivity (more than 0.25 s) with information sharing between the PFC and category-specific areas, manifested in long-range gamma-band or beta-band synchronization[45].

On the basis of our preregistration (see Extended Data Table 1), IIT would be challenged in the absence of sustained content-specific synchronization between face–object selective areas and V1/V2; whereas a challenge for GNWT would be a lack of phasic connectivity (0.3–0.5 s) between category-selective areas and PFC. Given the dynamical nature of these predictions, iEEG and MEG provided the most informative empirical test. These predictions were tested by computing pairwise phase consistency (PPC)[46] between each category-selective time series (face-selective and object-selective nodes) and either the V1/V2 or the PFC time series in the intermediate (1.0 s) and long-stimulus-duration (1.5 s), task-irrelevant trials (see section 7.1.2 in Supplementary Information for task-relevant trials). Gamma activity was analysed because of its close link to neuronal spiking[47], which IIT considers a constituent property of the physical substrate of consciousness[5].

For iEEG, we analysed connectivity between electrodes showing face and object selectivity, using a different subset of electrodes to test connectivity with V1/V2 and PFC (see Methods; Extended Data Fig. 5a for ROIs and examples of face-selective and object-selective electrodes). Given the sparse electrode coverage, the requirement to focus on 'activated' electrodes (see Methods) was relaxed, although restricting it to only activated electrodes yielded similar results. We found increased category-selective synchrony between category-selective and V1/V2 electrodes (Extended Data Fig. 5b). These effects were early and brief (for example, less than 0.75 s), and restricted to low frequencies (2–25 Hz). This synchrony was mostly explained by the stimulus-evoked response (Extended Data Fig. 6a). These results fail to align with IIT's predictions: the activity was neither sustained nor observed in the gamma frequency band. Bayesian analysis further supported the null hypothesis ($BF_{01} = 1.15–4.9$). No content-selective PPC was found between face-selective and object-selective electrodes and PFC electrodes in the relevant time window, contradicting the prediction of GNWT (Extended Data Fig. 6a; $BF_{01} = 2.62–5.32$).

For MEG, we found selective synchronization between face-selective areas and both V1/V2 and PFC. These effects were again early, restricted to low frequencies (2–25 Hz), and mostly explained by stimulus-evoked responses (Extended Data Figs. 5d and 6b). Bayesian analysis of the gamma-band synchronization further supported the null hypothesis (all $BF_{01} > 3$).

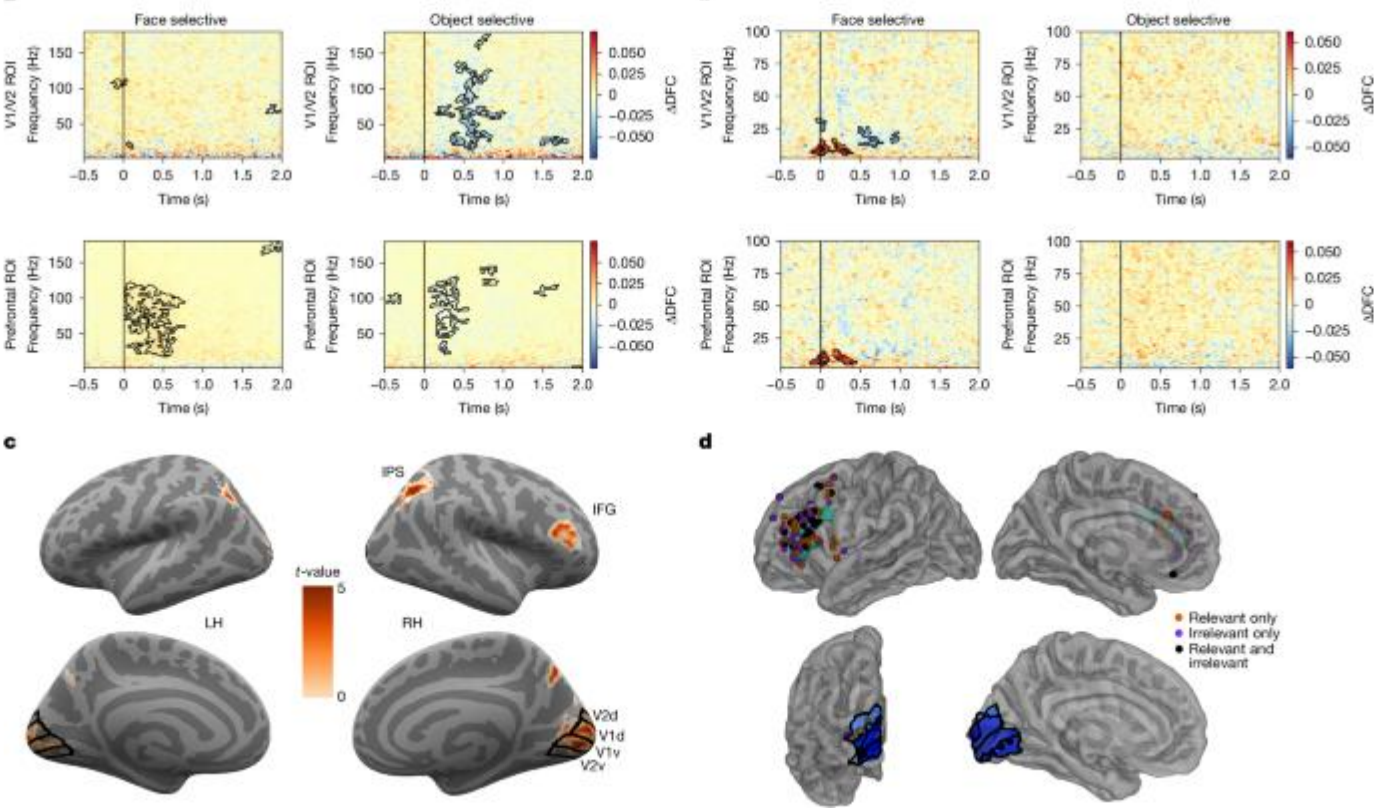
The results of the preregistered PPC metric for prediction 3, critical for both IIT and the GNWT, supported neither theory. PPC was chosen based on the mechanistic considerations of the theories because it assesses oscillatory phase. However, phase estimation from macroscopic recordings is susceptible to noise.

We thus used dynamic functional connectivity (DFC; see Methods), a metric sensitive to co-modulations of signal amplitude, after removing stimulus-evoked responses (Extended Data Fig. 6c,d includes the evoked response).

In iEEG, we observed significant connectivity between object-selective electrodes and V1/V2 (Fig. 4a), spanning multiple frequency bands, with the gamma band being the most predominant. In contrast to the predictions of IIT, the observed connectivity was brief. Connectivity between face-selective electrodes and V1/V2 was scarce, further supported by Bayesian analysis

($BF_{01} = 1.3$ in favour of the null hypothesis). Significant connectivity was observed between PFC and both the face-selective and the object-selective areas in the gamma frequency band within the GNWT-predicted time window. For MEG, brief DFC in the alpha–beta frequency bands was found between face-selective nodes and both PFC and V1/V2 (Fig. 4b).

Fig. 4: Prediction 3: interareal connectivity.



a, iEEG DFC analysis of task-irrelevant trials revealed significant content-selective synchrony only for object-selective electrodes in V1/V2 (for example, top row; face-selective: $n_{\text{patients}} = 4$ and $n_{\text{electrodes}} = 30$; object-selective: $n_{\text{patients}} = 4$ and $n_{\text{electrodes}} = 21$), while showing significant content-selective synchrony for both categories in the PFC ROI (bottom row; face selective: $n_{\text{patients}} = 19$ and $n_{\text{electrodes}} = 81$; object selective: $n_{\text{patients}} = 14$ and $n_{\text{electrodes}} = 57$). Here and in panel **b**, significance was assessed using a cluster-based permutation analysis ($P < 0.05$, two-sided) and the colour bars represent the average change in the DFC between conditions. **b**, MEG DFC analysis of task-irrelevant trials ($n = 65$) revealed significant content-selective synchrony below 25 Hz for the face-selective GED filter in both V1/V2 (top left) and PFC (bottom left), but not for the object-selective GED filter (right panels). **c**, fMRI generalized psychophysiological interaction (gPPI; $n = 70$) on task-relevant and task-irrelevant trials combined revealed significant content-selective connectivity when FFA is used as the analysis seed. Various significant regions showing task-related connectivity with the FFA seed were observed including V1/V2, right intraparietal sulcus (IPS) and right inferior frontal gyrus (IFG). LH, left hemisphere; RH, right hemisphere. **d**, Analysis of iEEG face-selective DFC synchrony across tasks is shown at the single-electrode level in PFC (top) and V1/V2 (bottom) ROIs. Electrodes showing significant synchrony (tested using a permutation test, FDR-corrected, $P < 0.05$) in relevant (orange-red), irrelevant (purple) or combined relevant and irrelevant (black) trials are shown (averaged over

70–120 Hz and 0–0.5-s time window). DFC synchrony was observed in both tasks, but restricted to IFG for the GNWT analysis and V2 regions for the IIT analysis, consistent with the fMRI gPPI analysis shown in panel **c**. Brain surfaces in panels **c**,**d** are from Freesurfer.

The exploratory DFC results in iEEG were consistent with the predictions of GNWT while challenging the predictions of IIT, as connectivity with V1/V2 was not sustained. However, V1/V2 were sparsely sampled in our population, with only 12 electrodes localized to V1/V2 compared with 472 in PFC.

Using fMRI, we evaluated connectivity across the entire cortex with homogeneous sampling. We computed the generalized psychophysiological interaction, defining the fusiform face area (FFA) and lateral occipital complex as seed regions (see Methods). Task-relevant and task-irrelevant trials were pooled to increase statistical power (see sections 7.1.1 and 12 in Supplementary Information for separate preregistered analyses). FFA showed content selective (face > object stimuli) connectivity with V1/V2, inferior frontal gyrus and intraparietal sulcus, consistent with predictions from both theories (Fig. 4c). No selective increase in interareal connectivity between object-selective nodes and PFC or V1/V2 was found in fMRI, even when separating task conditions (Extended Data Fig. 6f). Bayesian testing across prefrontal ROIs confirmed our findings with 62–94% of voxels across ROIs showing substantial evidence for the null hypothesis of no interareal connectivity ($BF_{01} = 3–7.75$). Support for the alternative hypothesis was observed only in 0–4% of voxels. The remaining voxels showed inconclusive evidence.

To determine whether connectivity to PFC and V1/V2 was task driven in the generalized psychophysiological interaction, we explored the iEEG data, separating task-relevant and irrelevant trials. We found task-independent, selective DFC connectivity (face > objects) for face-selective electrodes in both inferior frontal gyrus and V1/V2 (Fig. 4d).

Discussion

This adversarial collaboration aimed to address confirmation biases by researchers, breaking theoretical echo chambers[3] and identifying the strengths and weaknesses of theories[2,48] by forcing them to be explicit and committal about their empirical predictions, rigorously testing them on common methodological grounds[22,49], and providing the means for theorists to change their minds given conflicting results[49]. Doing so catalyses our ability to evaluate and arbitrate between theories of consciousness. Embracing this spirit, and adhering to guidelines for structuring adversarial collaborations[21], we opted for a three-voice discussion format, acknowledging that despite stringent testing of incompatible theoretical views, different interpretations of the same evidence may persist. Below, the theory-neutral consortium presents the main challenges our findings pose to the theories, based on the preregistered predictions, methods and analysis agreed on in advance with the adversaries. Then, adversaries offer their own interpretation of the findings and future directions (see sections 12 and 13 in Supplementary Information).

Extended Data Fig. 7 summarizes the key results, including the criteria used to assess whether findings support or contradict the theories. This summary covers both central and peripheral findings related to theory evaluation. The consortium adopted Lakatos’ sophisticated falsificationist approach to philosophy of science[13,50], emphasizing that challenged predictions provide more valuable insights than those confirmed by the data (see section 11 in Supplementary Information). Outcomes are weighted differentially across predictions and with respect to the different brain imaging modalities (Extended Data Table 1). This approach ensures a nuanced evaluation of the theories, highlighting areas of strength and those requiring further refinement.

For IIT, the lack of sustained synchronization within posterior cortex represents the most direct challenge, based on our preregistration. This is incompatible with IIT's claim that the state of the neural network, including its activity and connectivity, specifies the degree and content of consciousness[5]. Although this null result could stem from methodological limitations (for example, limited iEEG sampling of V1/V2 areas), our multimodal and highly powered study provided the best conditions so far for evaluating the prediction. We urge IIT proponents to direct future efforts to evaluate this prediction and determine the implications of this failure.

More broadly, although IIT passed the predefined criteria for the duration prediction (number 2), there was no evidence for a sustained representation of orientation, despite being a property of the consciously perceived stimuli[25]. This is an informative challenge for IIT, as orientation decoding was robust across all three data modalities, leaving open the question of whether and how information about orientation is maintained over time.

For GNWT, the most substantial challenge based on our preregistered criteria pertains to its account for the maintenance of a conscious percept over time and, in particular, the lack of ignition at stimulus offset. This result is unlikely to stem from sensitivity limitations, as offset responses were robustly found elsewhere (for example, visual areas); and in PFC, strong onset responses were found to the very same stimuli. The lack of ignition at stimulus offset is especially surprising given the change in conscious experience at the onset of the blank fixation screen. This clear update to the content of consciousness should have been represented somehow by the global workspace[12]. Thus, that aspect of consciousness remains unexplained within the GNWT framework.

Another key challenge for GNWT pertains to representing the contents of experience: although we found representation of category in PFC irrespective of the task, thereby demonstrating the sensitivity of our methods, no representation of identity was found, and representation of orientation was evident only in MEG (signal leakage notwithstanding); although these dimensions were clearly a part of the conscious experience of participants of the stimuli. This raises the question of whether PFC is involved in broadcasting all conscious content, as predicted by GNWT[23], or only a subset (for example, abstract concepts and categories, rather than low-level details), in which case the role of PFC in consciousness might need to be redefined.

Before this study, predictions from IIT and GNWT were typically tested with one data modality at a time[23,24], leaving room for negative results to be easily attributed to the limitations of the chosen modality[51]. We combined multiple techniques (iEEG, MEG and fMRI) to mitigate these limitations, cross-compensating for their weaknesses. This methodological approach was mutually agreed upon by the theory leaders before data collection and results disclosure as the most powerful and conclusive approach, making both positive and negative findings more meaningful.

Although this study was designed around IIT and GNWT, the results may have implications for other theories of consciousness. For example, the prediction of GNWT about PFC is shared by those higher-order theories of consciousness that hypothesize that PFC actually supplies the content of visual consciousness (for example, ref. [17]), rather than those that take it to merely enable the consciousness of content that is located in posterior visual areas (for example, refs. [52,53]). As a result, the failures to confirm this prediction challenge not only GNWT but also those higher-order theories[54]. Predictions 2 and 3, about timing and connectivity, are more distinctive to GNWT but could also be shared by other theories. Likewise, the non-core prediction 1 about the posterior cortex by IIT is also shared by many theories (for example, recurrent processing theory[14]), and its prediction 2 about timing may be shared by some posterior theories of consciousness, such as the local recurrency theory[15]. Its prediction 3 about interareal connectivity is more distinctive to IIT (for example, it is not shared by synchrony theory[55]), so the challenge here is more specific as well.

Our study focused on the contents of consciousness (for example, category, identity, orientation and duration), linking brain activity to subjective phenomenology. This departs from the traditional contrastive method, which compares the presence and absence of consciousness but conflates it with other cognitive processes (for example, decision-making or memory formation)[56,57,58]. Some might argue that our approach tracks stimulus processing rather than consciousness. Yet, our aim is to challenge and potentially falsify[50,59] IIT and GNWT, by examining where their predictions differ, rather than to discover the neural correlates of consciousness. In this context, what might seem like a weakness—focusing on the presence of fully attended, consciously experienced stimuli—is actually beneficial for testing the primary positive predictions of the theories and their failures. This is because such failures are harder to dismiss owing to weak signals. Thus, our approach assesses whether the proposed neural mechanisms are truly necessary for consciousness.

Our study, although comprehensive, is not without limitations. First, we cannot entirely rule out residual task engagement with respect to category, although our design ensured that orientation and duration remained task-irrelevant, so the results on these dimensions cannot be explained by task-related effects. Second, although we aimed to capture multiple aspects of consciousness, our approach still falls short of encompassing its full phenomenal richness. Third, despite the high spatial and temporal resolution of our data, it lacks single-unit recordings, which are typically restricted to patients with epilepsy and selected brain regions. Ongoing studies in animal models, as part of a separate adversarial collaboration, can accordingly complement our findings.

Beyond directly challenging the theories, our study raises broader questions about theory testing and development across disciplines. A key challenge is how to weigh predictions and integrate evidence across different analyses and measurement techniques (for example, fMRI, MEG and iEEG). We adopted a lenient falsificationist approach, in which evidence for any tested feature (for example, decoding of category or orientation) was sufficient to rule out failure, rather than requiring consistency across all features. However, a formal framework is urgently needed to quantitatively integrate evidence, accounting for prediction centrality, measurement error and cross-sample consistency. Such a framework would enhance systematic theory building in an era of accumulating results[60].

After reviewing the results and the discussions by adversaries, readers might expect a definitive verdict on the two theories under evaluation. Instead, we invite readers to weigh the evidence themselves—considering the support for each preregistered prediction, the breadth of the data, the sophistication of the methods and analyses, and the cognitive biases that shape interpretation. Scientific progress is rarely a matter of simple verdicts; evidence is filtered through previous beliefs and motivations[61], making theory evaluation a dynamic process. By presenting results and adversarial responses transparently, we embrace the openness needed for science to converge on robust explanations of complex phenomena such as consciousness.

Methods

Ethics statement

The experiment was approved by the institutional ethics committees of each participating data-collecting laboratory, including the Science, Technology, Engineering and Mathematics Ethical Review Committee at the Centre for Human Brain Research, University of Birmingham (ERN_18-0226AP20); the Committee for Protecting Human and Animal Subjects at the School of Psychological and Cognitive Sciences, Peking University (2020-05-07e); the Commissie Mensgebonden Onderzoek Regio Arnhem-Nijmegen at the Centre for Cognitive Neuroimaging at Donders Institute (NL45659.091.14); the Human Research Protection Program Institutional Review Board at Yale School of Medicine (2000027591); the Office of Science and Research Institutional Review Board at New York University Langone Health (i14-02101_CR6); the Boston Children’s Hospital Institutional Review Board at Children’s Hospital Corporation d/b/a Boston Children’s Hospital (04-05-065R); the Institutional Review Board at the University of Wisconsin-Madison (ID: 2017-1299); and the Ethics Council of the Max Planck Society at Max Planck Institute for Empirical Aesthetics (Nr. 2017_12). All participants and patients provided oral and written informed consent before participating in the study. All study procedures were carried out in accordance with the Declaration of Helsinki. Patients were also informed that clinical care was not affected by participation in the study.

Participants

Healthy participants and patients with pharmaco-resistant focal epilepsy participated in this study. The datasets reported here consist of: (1) behaviour, eye tracking and iEEG data collected at the Comprehensive Epilepsy Center at New York University (NYU) Langone Health, the Brigham and Women’s Hospital, the Boston Children’s Hospital (Harvard), and the University of Wisconsin School of Medicine and Public Health (WU). (2) Behaviour, eye tracking, MEG and EEG data collected at the Centre for Human Brain Health (CHBH) of the University of Birmingham (UB), and at the Center for MRI Research of Peking University (PKU). (3) Behaviour, eye tracking and fMRI data collected at the Yale Magnetic Resonance Research Center (MRRC) and at the Donders Centre for Cognitive Neuroimaging (DCCN), of Radboud University Nijmegen. For both the MEG and fMRI datasets, one-third of the data that passed quality tests (henceforth, the optimization dataset; see the section 'Preregistration' for details about quality test criteria[27]) were used to optimize the analysis methods, which were subsequently added to the preregistration as an additional amendment. These preregistered analyses were then run on the remaining two-thirds of the data (henceforth, the replication dataset) and constitute the data reported in the main study. This procedure was not used for the iEEG data due to the serendipitous nature of the recording and electrode placement, the rarity of this type of data and the increased difficulty of data collection due to the COVID-19 pandemic.

A total of 97 healthy participants were included in the MEG sample (mean age of 22.79 ± 3.59 years, 54 females, all right handed), 32 of those datasets were included in the optimization phase (mean age of 22.50 ± 3.43 years, 19 females, all right handed), and 65 in the replication sample (mean age of 22.93 ± 3.66, 35 females, all right handed). Five additional participants were excluded from the MEG dataset: two because of failure to meet predefined behavioural criteria (that is, Hits of less than 80% and/or False Alarms > 20%), two because of excessive noise from sensors, and one because of incorrect sensor reconstruction. A total of 108 healthy participants were included in the fMRI sample (mean age of 23.28 ± 3.46 years, 70 females, 105 right handed); 35 of those datasets were included in the optimization sample (mean age of 23.26 ± 3.64 years, 21 females, 34 right handed) and 73 in the replication sample (mean age of 23.29 ± 3.37, 49 females, 71 right handed). Twelve additional participants were excluded from the fMRI dataset: eight because of motion artefacts, two because of insufficient coverage and two because of incomplete data (with respect to these last two participants, see section 14 of the Supplementary Information for deviations from the preregistration document). For the iEEG arm of the project, a total of 34 patients were recruited. Two patients were excluded owing to incomplete data. Demographic, medical and neuropsychological scores for each patient, when available, are reported in Supplementary Table 25. Three iEEG patients whose behaviour fell slightly short of the predefined behavioural criteria (that is Hits of less than 70%, FA > 30%) were nonetheless included given the difficulty of obtaining additional iEEG data (see section 14 in Supplementary Information for deviation from the preregistration).

Experimental procedure

Experimental design

To test critical predictions of the theories, five experimental manipulations were included in the experimental design: (1) four stimulus categories (faces, objects, letters and false fonts), (2) 20 stimulus identities (20 different exemplars per stimulus category), (3) three stimulus orientations (front, left and right view), (4) three stimulus durations (0.5 s, 1.0 s and 1.5 s), and (5) task relevance (relevant targets, relevant non-targets and irrelevant).

Stimulus category, stimulus identity and stimulus orientation served to test predictions about the representation of the content of consciousness in different brain areas by the theories. In addition, stimulus duration served to test predictions about the temporal dynamics of sustained conscious percepts and interareal synchronization between areas. Task relevance served to rule out the effect of task demands, as opposed to conscious perception per se, on the observed effects[62]. This aspect of the experimental design was inspired by ref. [63].

Stimuli

Four stimulus categories were used: faces, objects, letters and false fonts. These stimuli naturally fell into two clearly distinct groups: pictures (faces and objects) and symbols (letters and false fonts). These natural couplings were aimed at creating a clear difference between task-relevant and task-irrelevant stimuli in each trial block (see the section ‘Procedure’). All stimuli covered a squared aperture at an average visual angle of 6° by 6°. Face stimuli were created with FaceGen Modeler 3.1; letter and false font stimuli were generated with MAXON CINEMA 4D Studio (RC - R20) 20.059; object stimuli were taken from the Object Databank[64]. Stimuli were grey scaled and equated for luminance and size. To facilitate face individuation, faces had different hairstyles and belonged to different ethnicities and genders. Equal proportions of male and female faces were presented. The orientation of the stimuli was manipulated, such that half of the stimuli from each category had a side view (30° and −30° horizontal viewing angle, left and right orientation) and the other half had a front view (0°).

Procedure

Participants performed a non-speeded target detection task (see Supplementary Video 1). The experiment was divided into runs, with four blocks in each run (see the section ‘Trial counts’). On a given block, participants viewed a sequence of single, supra-threshold, foveally presented stimuli belonging to one of four stimulus categories and presented for one of three stimulus durations onto a fixation cross that was present throughout the experiment. Within each block, half of the stimuli were task-relevant and half were task-irrelevant. To manipulate task relevance, at the beginning of each block participants were instructed to detect the rare occurrences of two target stimulus identities, one from each relevant category (for pictures, face–object; for symbols, letter–false font), irrespective of their orientation. This was specified by presenting the instruction ‘detect face A and object B’ or ‘detect letter C and false font D’, accompanied by images for each target (see Fig. 1d). Targets did not repeat across blocks. Each run contained two blocks of the face–object task and two blocks of the letter–false font task, with block order counterbalanced across runs.

Accordingly, each block contained three different trial types: (1) targets: the two stimuli being detected (for example, the specific face and object identities); (2) task-relevant stimuli: all other stimuli from the task-relevant categories (for example, the non-target faces–objects); and (3) task-irrelevant stimuli: all stimuli from the two other categories (for example, letters–false fonts). An advantage of this design is that the three trial types enabled a differentiation of neural responses related to task goal, task relevance and simply consciously seeing a stimulus. We confirmed that participants were conscious of the stimuli in both the task-relevant and task-irrelevant trials in a separate experiment, which included a surprise memory test (see section 3 in Supplementary Information).

Stimuli were presented for one of three durations (0.5 s, 1.0 s or 1.5 s), followed by a blank period of a variable duration to complete an overall trial length fixed at 2.0 s. For the MEG and iEEG version, random jitter was added at the end of each trial (mean inter-trial interval of 0.4 s, jittered 0.2–2.0 s, truncated exponential distribution) to avoid periodic presentation of the stimuli. The mean trial length was 2.4 s. For the fMRI protocol, timing was adjusted as follows: the random jitter between trials was increased (mean inter-trial interval of 3 s, jittered 2.5–10 s, with truncated exponential distribution), with each trial lasting approximately 5.5 s. This modification helped with avoiding non-linearities in BOLD signal, which may affect fMRI decoding[65]. Second, to increase detection efficacy for amplitude-based analyses, three additional baseline periods (blank screen) of 12 s each were included per run (total of 24). The identity of the stimuli was randomized with the constraint that they appeared equally across durations and tasks conditions. Participants were further instructed to maintain central fixation on a black circle with a white cross and another black circle in the middle throughout each trial (see Supplementary Fig. 1d and Supplementary Video 1 for a demonstration of the experimental paradigm).

Trial counts

The MEG study consisted of 10 runs containing 4 blocks each with 34–38 trials per block, 32 non-targets (8 per category) and 2–6 targets, for a total of 1,440 trials. The same design was used for iEEG, but with half the runs (5 runs total), resulting in a total of 720 trials. For fMRI, there were 8 runs containing 4 blocks each with 17–19 trials per block, 16 non-targets (4 per category) and 1–3 targets, for a total of 576 trials. Rest breaks between runs and blocks were included.

Data acquisition

Behavioural data acquisition

The task was run on Matlab (PKU: R2018b; DCCN, UB and Yale: R2019b; Harvard: R2020b; NYU: R2020a, and WU: 2021a) using Psychtoolbox (v3)[66]. The iEEG version of the task was run on a Dell Precision 5540 laptop, with a 15.6″ Ultrasharp screen at NYU and Harvard and on a Dell D29M PC with an Acer 19.1″ screen in WU. Participants responded using an eight-button response box (Millikey LH-8; response hand (or hands) varied based on the setting in the patient's room). The MEG version was run on a custom PC at UB and a Dell XPS desktop PC on PKU. Stimuli were displayed on a screen placed in front of the participants with a PROPixx DLP LED projector (VPixx Technologies). Participants responded with both hands using two 5-button response boxes (NAtA or SINORAD). The fMRI version was run on an MSI laptop at Yale and a Dell Desktop PC at DCCN. In DCCN, stimuli were presented on an MRI compatible Cambridge Research Systems BOLD screen 32″ IPS LCD monitor, and in Yale they were presented on a Psychology Software Tools Hyperion projection system to project stimuli on the mirror fixed to the head coil. Participants responded with their right hand using a 2 × 2 current designs response box at Yale and a 1 × 4 current designs response box at DCCN.

Eye tracking data acquisition

For the iEEG setup, eye tracking and pupillometry data were collected using a EyeLink 1000 Plus in remote mode, sampled monocularly at 500 Hz (from the left eye at WU, and depending on the setup at Harvard), or on a Tobii-4C eye tracker, sampled binocularly at 90 Hz (NYU). The MEG and fMRI laboratories used the MEG-compatible and fMRI-compatible EyeLink 1000 Plus Eye-tracker system (SR Research) to collect data at 1,000 Hz. For MEG, eye tracking data were acquired binocularly. For fMRI, data were acquired monocularly from either the left or the right eye, in DCCN and Yale, respectively. For all recordings, a 9-point calibration was performed (besides Harvard, where a 13-point calibration was used) at the beginning of the experiment, and recalibration was carried out as needed at the beginning of each block or run.

iEEG data acquisition

Brain activity was recorded with a combination of intracranial subdural platinum-iridium electrodes embedded in SILASTIC sheets (2.3-mm diameter contacts, Ad-Tech Medical Instrument and PMT Corporation) and/or depth stereo-electroencephalographic platinum-iridium electrodes (PMT Corporation; 0.8 mm in diameter, 2.0-mm length cylinders; separated from adjacent contacts by 1.5–2.43 mm), or depth stereo-electroencephalographic platinum-iridium electrodes (BF08R-SP21X-0C2, Ad-Tech Medical; 1.28 mm in diameter, 1.57 mm in length, 3–5.5-mm spacing). Electrodes were arranged as grid arrays (either 8 × 8 with 10-mm centre-to-centre spacing, 8 × 16 contacts with 3-mm spacing, or hybrid macro–micro 8 × 8 contacts with 10-mm spacing and 64 integrated microcontacts with 5-mm spacing), linear strips (1 × 8/12 contacts), depth electrodes (1 × 8/12 contacts) or a combination thereof. Recordings from grid, strip and depth electrode arrays were done using a Natus Quantum amplifier or a Neuralynx Atlas amplifier. A total of 4,057 electrodes (892 grids, 346 strips and 2,819 depths) were implanted across 32 patients with drug-resistant focal epilepsy undergoing clinically motivated invasive monitoring. A total of 3,512 electrodes (780 grids, 307 strips and 2,425 depths) that were unaffected by epileptic activity, artefacts or electrical noise were used in subsequent analyses. To determine the electrode localization for each patient, a post-operative computed tomography scan and a pre-operative T1 MRI were acquired and co-registered.

MEG data acquisition

MEG was acquired using a 306-sensor TRIUX MEGIN system, comprising 204 planar gradiometres and 102 magnetometres in a helmet-shaped array. The MEG gantry was positioned at 68° for optimal coverage of frontal and posterior brain areas. Simultaneous EEG was recorded using an integrated EEG system and a 64-channel electrode cap (EEG data are not reported here, but are included in the shared dataset). During acquisition, MEG and EEG data were bandpass filtered (0.01 and 330 Hz) and sampled at 1,000 Hz. The location of the head fiducials, the shape of the head, the positions of the 64 EEG electrodes and the head position indicator (HPI) coil locations relative to anatomical landmarks were collected with a 3D digitizer system (Polhemus Isotrack). ECG was recorded with a set of bipolar electrodes placed on the chest of the participant. Two sets of bipolar electrodes were placed around the eyes (two at the outer canthi of the right and left eyes and two above and below the centre of the right eye) to record eye movements and blinks (EOG). Ground and reference electrodes were placed on the back of the neck and on the right cheek, respectively. The head position of participants on the MEG system was measured at the beginning and end of each run, and also before and after each resting period, using four HPI coils placed on the EEG cap, next to the left and right mastoids and over left and right frontal areas.

Anatomical MRI data acquisition

For source localization of the MEG data with individual realistic head modelling, a high-resolution T1-weighted MRI volume (3 T Siemens MRI Prisma scanner) was acquired per participant. Anatomical scans were acquired either with a 32-channel coil (repetition time (TR)/echo time (TE) = 2,000/2.03 ms; inversion time (TI) = 880 ms; 8° flip angle; field of view = 256 × 256 × 208 mm; 208 slices; 1-mm isotropic voxels, UB) or a 64-channel coil (TR/TE = 2,530/2.98 ms; TI = 1,100 ms; 7° flip angle; field of view = 224 × 256 × 192 mm, 192 slice, 0.5 × 0.5 × 1 mm voxels, PKU). The FreeSurfer standard template was used (fsaverage) for participants lacking an anatomical scan ($n = 5$).

fMRI data acquisition

MRI data were acquired using a 32-channel head coil on a 3 T Prisma scanner. A session included high-resolution anatomical T1-weighted MPRAGE images (GRAPPA acceleration factor = 2, TR/TE = 2,300/3.03 ms, 8° flip angle, 192 slices, 1-mm isotropic voxels), and a whole-brain T2*-weighted multiband-4 sequence (TR/TE = 1,500/39.6 ms, 75° flip angle, 68 slices, voxel size of 2 mm isotropic, anterior/posterior (A/P) phase-encoding direction, field of view = 210 mm, bandwith (BW) = 2,090 Hz px^{-1}). A single-band reference image was acquired before each run. To correct for susceptibility distortions, additional scans using the same T2*-weighted sequence, but with inverted phase-encoding direction (inverted readout/phase-encoding (RO/PE) polarity), were collected while the participant was resting at multiple points throughout the experiment.

Preprocessing and analysis details

For readability, we first detail the preprocessing protocols for each of the modalities (iEEG, MEG and fMRI) separately. Then, we describe the different analyses, combining information across the modalities, while noting any differences between them.

Analysis strategy

As part of our testing framework, after excluding a limited number of participants due to data quality checks, we conducted an initial optimization phase on one-third of the MEG ($n = 32$) and fMRI ($n = 35$) datasets to evaluate data quality across sites and to optimize analysis pipelines. Following the optimization phase, pipelines were preregistered[27] and applied to the novel datasets containing twice as much data (MEG $n = 65$ and fMRI $n = 73$).

In the main paper, we report results obtained on the novel, previously unexamined datasets. For iEEG, given the smaller sample, a different analysis strategy was implemented. We refer the reader to the iEEG methods section and text in the main paper for numbers of participants that were entered in each analysis. Results from the optimization phase are reported in section 4 of Supplementary Information. The results of the optimization phase and the preregistered replication phase were compared and deemed to be largely compatible, with some minor exceptions (section 4 of Supplementary Information).

iEEG preprocessing

Data were converted to BIDS[67] and preprocessed using MNE-Python (v0.24)[68], and custom-written functions in Python and Matlab. Preprocessing steps included downsampling to 512 Hz, detrending, bad channel rejection, line noise and harmonic removal, and re-referencing. Electrodes were re-referenced to a Laplacian scheme[69], whereas bipolar referencing was used for electrodes at the edge of a strip, grid or stereo EEG, and the signal was localized at the midpoint (Euclidean distance) between the two electrodes. Electrodes with no direct neighbours were discarded. Seizure-onset zone electrodes, those localized outside the brain and/or containing no signal or high amplitude noise level were discarded. Line noise and harmonics were removed using a one-pass, zero-phase non-causal band-stop FIR filter.

The high-gamma power (70–150 Hz) was obtained by bandpass filtering the raw signal in eight successive 10-Hz-wide frequency bands, computing the envelope using a standard Hilbert transform, and normalizing it (dividing) by the mean power per frequency band across the entire recording. To produce a single high-gamma envelope time series, all frequency bands were averaged together[70]. Most analyses focused on the high-gamma power as it closely correlated with neural spiking activity[71] and with the BOLD signal[37]. To obtain the event-related potentials (ERPs), the raw signal was low-pass-filtered at 30 Hz with a one-pass, zero-phase, non-causal low-pass FIR filter. Epochs were segmented between 1-s pre-stimulus until 2.5-s post-stimulus of interest.

Surface reconstruction and electrode localization

Electrode positions were determined based on a computed tomography scan coregistered with a pre-implant T1-weighted MRI. A 3D reconstruction of the brain of each patient was computed using FreeSurfer (http://surfer.nmr.mgh.harvard.edu). For visualization, the electrode positions for individual participants were converted to the Montreal Neurological Institute (MNI)152 space. As each theory specified a set of anatomical ROIs, after electrode localization, electrodes were labelled according to the Freesurfer-based Destrieux atlas segmentation[72,73] and/or Wang atlas segmentation[74].

Identification of task-responsive channels

To identify task-responsive electrodes, we computed the area under the curve (AUC) for the baseline (−0.3 to 0 s) and the stimulus-evoked period (0.05–0.35 s) separately for the task-relevant and task-irrelevant conditions, and compared them per electrode using a Wilcoxon sign-rank test, corrected for false discovery rate (FDR)[75]. A Bayesian *t*-test[76] was used to quantify evidence for non-responsiveness.

Identification of category-selective channels

To determine category selectivity for faces, objects, letters and false fonts in the high gamma, we followed the method of Kadipasaoglu and colleagues[77]. Per category, we computed a *d′* (AUC of 0.05–0.4 s) comparing the activation between the category of interest (u_j) and each of the other categories (u_i), normalized by the standard deviation of each category:

d′=uj−1N∑iNui12(σj2+1N∑iNσi2);i≠j

A permutation test (10,000 permutations) was used to evaluate significance. *d′* was computed for the task-relevant and task-irrelevant conditions separately. An electrode was considered selective if it showed selectivity on both tasks.

Multivariate analysis electrodes combination

Owing to the sparse and highly variable coverage of iEEG data, all collected electrodes were combined into a ‘super participant’ multivariate analyses (RSA and decoding). To create a single-trial matrix for the super participant, we equated the trial matrices of all our participants by subsampling to the lowest number of trials in the relevant conditions. Participants that did not complete the full experiment were discarded ($n = 3$), resulting in a total of 29 participants with 583 electrodes in posterior ROIs and 576 electrodes in prefrontal ROIs. For analyses on stimuli identities, stimuli that were presented less than three times to any of the participants across intermediate and long trials in the task-relevant and task-irrelevant trials were discarded. We then subsampled the trials for each identity to three trials per participant. The subsampling procedure was repeated 100 times to avoid random fluctuations induced by the subsampling. The analysis was computed for each repetition and average across repetitions.

MEG preprocessing

The MEG data were converted to BIDS[78] using MNE-BIDS[79], and preprocessed following the FLUX Pipeline[80] in MNE-Python (v0.24.0)[68]. Preprocessing steps included MEG sensor reconstruction using a semi-automatic detection algorithm and signal-space separation[81] to reduce environmental artefacts. FastICA[82] was used to detect and remove cardiac and ocular components from the data for each participant (mean = 2.90 components, s.d. = 0.92). Before ICA, data were segmented, and segments containing muscle artefacts were removed. After preprocessing, data were epoched into 3.5-s segments (1-s pre-stimulus to 2.5-s post-stimulus onset). Trials in which gradiometre values exceeded 5,000 fT cm^{-1}, magnetometres exceeded 5,000 fT and/or the trial contained muscle artefacts were rejected from the MEG dataset. Finally, to be included in the analyses, participants should have a minimum of 30 clean trials per condition. No participants were excluded because of not meeting this criterion.

Source modelling

MEG source modelling was performed using the dynamic statistical parametric mapping method[83], based on depth-weighted minimum-norm estimates (MNEs)[84,85], on epoched and baseline (−0.5 s to 0 s before stimulus onset) corrected data. To build a forward model, the MRI images were manually aligned to the digitized head shape. A single shell boundary elements model was constructed in MNE-Python based on the inner skull surface derived from FreeSurfer[72,73], to create a volumetric forwards model (5-mm grid) covering the full-brain volume. The lead field matrix was then calculated according to the head position with respect to the MEG sensor array. A noise covariance matrix for the baseline and a covariance matrix for the active time window were calculated and the combined (that is, sum) covariance matrix was used with the forwards model to create a common spatial filter. Data were spatially pre-whitened using the covariance matrix from the baseline interval to combine gradiometre and magnetometre data[86].

fMRI preprocessing

Source DICOM data were converted to BIDS using BIDScoin (v3.6.3)[87]. This includes converting DICOM data to NIfTI using dcm2niix[88] and creating event files using custom Python codes. BIDS compliance of the resulting dataset was controlled using BIDS-Validator. Subsequently, MRI data quality control was performed using MRIQC (0.16.1)[89] and custom scripts for data rejection. All (f)MRI data were preprocessed using fMRIPrep (20.2.3)[90], based on Nipype (1.6.1)[91]. For further details on the fMRIprep pipeline, see preregistration. Custom scripts used NumPy (1.19.2)[92] and Pandas (1.1.3)[93].

Analysis-specific functional preprocessing

Additional, analysis-specific, fMRI data preprocessing was performed using FSL 6.0.2 (FMRIB Software Library)[94], Statistical Parametric Mapping (SPM 12) software[95], and custom Python scripts (using NiBabel (3.2.2)[96] and SciPy (1.8.0)[97] after the above-outlined general preprocessing. Functional data for univariate data analyses were spatially smoothed (Gaussian kernel with full-width at half-maximum of 5 mm), grand mean scaled and temporal high-pass filtered (128 s). No spatial smoothing was applied for multivariate analyses.

Contrast of parameter estimates

We modelled BOLD signal responses to the experimental variables by fitting voxel-wise general linear model (GLM) to the data of each run using FSL FEAT. The following regressors were modelled in an event-related approach, with event duration corresponding to the stimulus duration (that is, 0.5 s, 1.0 s and 1.5 s), and convolved with a double gamma haemodynamic response function: 12 regressors of interest (targets, task-relevant and task-irrelevant stimuli per stimulus category, that is, faces, objects, letters and false fonts; and a regressors of no interest, that is, target screen display). We included the first-order temporal derivatives of the regressors of interest, and a set of nuisance regressors: 24 motion regressors (FMRIB Software Library (FSL)'s standard + extended set of motion parameters) plus a cerebrospinal fluid (CSF) and a white matter (WM) tissue regressor. Each of the 12 regressors of interest was contrasted against an implicit baseline (used in the putative Neural Correlates of Consciousness analysis; see below). In addition, we obtained contrast of parameter estimates for 'relevant faces versus relevant objects', 'relevant letters versus relevant false fonts', 'irrelevant faces versus irrelevant objects', 'irrelevant letters versus irrelevant false fonts' (used for the definition of decoding ROIs), 'relevant and irrelevant faces versus relevant and irrelevant objects' and 'all stimuli versus baseline' (used for the definition of seeds for the generalized psychophysiological interaction (gPPI) analysis). Data were averaged across runs per participant using FSL's fixed-effects analysis and subsequently averaged across participants using FSL's FLAME1 mixed-effect analysis. Gaussian random-field cluster thresholding was used to correct for multiple comparisons, using the default settings of FSL, with a cluster formation threshold of one-sided $P < 0.001$ ($z \geq 3.1$) and a cluster significance threshold of $P < 0.05$.

Anatomical ROIs

ROIs were defined a priori in consultation with the adversarial theories. They were determined per participant based on the Destrieux atlas[73] including both hemispheres, and then resampled to standard MNI space (see Supplementary Table 26). For the connectivity analysis, areas V1/V2 (combining dorsal and ventral) were defined based on the Wang cortical parcellation[74]. For details on the process of selecting the ROIs and the justification of the ROI selection in the context of this study, see section 10 in Supplementary Information. All anatomical segmentations were performed using Freesurfer (6.0.1)[72].

Behavioural analyses

Log-linear-corrected d' (ref. [98]), false alarms and reaction times were computed per category and stimulus duration, separately (false alarms were also calculated per task relevance, without duration) and per modality (iEEG, MEG and fMRI). These measures were compared with linear–logistic mixed models, where appropriate. For the former, we report analysis of variance omnibus F-tests, and for the latter, omnibus χ^2 test from an analysis of deviance. We approximated degrees of freedom using the Satterthwaite method[99]. Pairwise t-tests following significant interactions were Bonferroni corrected. To estimate Bayesian information criterion (BIC) differences between the original and null logistic models, we used the P values and sample size[100] (p_to_bf package in R).

Eye-tracking analyses

For Eyelink, gaze and pupil data were segmented, and trials with missing data were excluded. Blinks were detected using the Hershman algorithm[101], and removed with 200-ms padding[102]. The Eyelink standard parser algorithm was used for saccade and fixation detection. Saccades were further corroborated using the Engbert and Kliegl[103] algorithm. Fixations were baseline corrected (−0.25 s to 0 s). Mean fixation distance, mean blink rate, mean saccade amplitude and mean pupil size were compared in a LMM with category and task relevance as fixed effects, and participant and item as random effects. Separate analyses were carried out on the first 0.5 s after stimulus onset including all trials; and on the 1.5-s trials including time window (0–0.5 s, 0.5–1.0 s and 1.0–1.5 s) as fixed effects. BIC was used to test the models against the null hypothesis models. For Tobii, gaze coordinate data were segmented, missing data were excluded and coordinates were baseline corrected to depict heatmaps of patients' gaze. Of note, the coordinate data were not added to the LMMs due to its poorer quality with respect to the EyeLink data.

Decoding analysis

All decoding analyses were performed using a linear support vector machine (SVM; scikit learn (0.23.2), https://scikit-learn.org/) classifier. Below, we explain how this was done for each one of the predictions.

iEEG decoding was done on the high-gamma signal, averaged over non-overlapping windows of 0.02 s separately for electrodes located in the GNWT and IIT ROIs. The top 200 electrodes (selectKbest[104]), as determined by a F-test within a given set of

electrodes from the theory ROIs, were used as features for the classifier. Two-hundred features were selected to provide a balance between model optimization (for example, feature selection) and participant representation (for example, electrodes or features coming from multiple participants). Statistical significance of decoding performance was assessed via permutation test, randomly permuting the sample labels and repeating the decoding analysis 1,000 times, corrected for multiple comparisons using a cluster-based correction (cluster mass inference with cluster forming threshold at $P < 0.05$)[105,106]. Also, to assess the decoding accuracy within unique ROIs (for example, S_temporal_sup of the Destrieux atlas), separate classifiers were trained using all electrodes in a given parcel. Each classifier was fitted using all electrodes in a parcel and time window (GNWT: 0.3–0.5 s, IIT: 0.3–1.5 s) as features, resulting in a single accuracy value per parcel. SelectKbest (200 features for iEEG) feature selection and fivefold cross-validation with three repetitions was used. To assess the statistical significance of the decoding accuracy within unique ROIs (so only one accuracy score is obtained per ROI), P values obtained via permutation tests were corrected for multiple comparisons across all ROIs using FDR correction ($q \leq 0.05$[75]). To compute Bayes factors on the decoding accuracy values, we used a β-binomial approach that compares the marginal likelihood under a point-null hypothesis against a flat $B(\alpha=1,\beta=1)$ alternative prior, yielding an analytic Bayes factor. We then derived the null hypothesis parameters from the empirical null distribution by updating a tight prior centred at chance level ($B(\alpha=1{,}000,\beta=1{,}000)$) with the shuffle-based accuracies, thereby incorporating any bias present in the null distribution.
MEG decoding was done on bandpass-filtered (1–40 Hz) and downsampled (100 Hz) data. The reconstructed source-level MEG data within a subset of the predefined anatomical ROIs (GNWT: 'G_and_S_cingul-Ant', 'G_and_S_cingul-Mid-Ant', 'G_and_S_cingul-Mid-Post', 'G_front_middle', 'S_front_inf', 'S_front_sup'; IIT: 'G_cuneus', 'G_oc-temp_lat-fusifor', 'G_oc-temp_med-Lingual', 'Pole_occipital', 'S_calcarine', 'S_oc_sup_and_transversal', as they showed high response to the stimulus on the optimization dataset) were extracted for further analysis (500 vertices and 800 vertices per hemisphere for each of the anatomical ROI defined by the theories). We applied temporal smoothing (0.05-s window, 0.01-s sliding window), computed pseudotrials[107], normalized the data and selected the top 30 features within a given ROI as features for the different classifiers. A group-level one-sample t-test per time point was performed on the decoding accuracy results, corrected for multiple comparisons using a cluster-based correction[106].

The overall decoding strategy for fMRI was similar to that used on the iEEG and MEG data, yet with some differences. A multivariate pattern analysis approach was used on the pattern of BOLD activity over voxels. A non-spatially smoothed parameter estimate map was obtained by fitting a GLM per event with that event as the regressor of interest and all the other remaining events as one regressor of no interest[108] as implemented in NiBetaSeries (0.6.0) package. The model also included the 24 nuisance regressors described in the 'fMRI preprocessing' section.

Decoding was performed using whole-brain and ROI-based approaches. The whole-brain analysis was performed using a searchlight approach with 4-mm radius. For ROI-based decoding, decoding ROIs were defined based on functional fMRI contrasts (see the 'fMRI preprocessing' section) and constrained with pre-defined anatomical ROIs (see Extended Data Table 2 on anatomical ROIs). A one-sample permutation test was used to determine whether decoding significantly exceeded chance level within each ROI. FDR was used to correct for multiple comparisons across ROIs. For whole-brain decoding, a cluster-based permutation test was used to evaluate the decoding statistical significance across participants ($P < 0.05$), complemented by Bayesian analysis. In addition, stimulus versus baseline searchlight decoding was performed using leave-one-run out cross-validation, and the resultant decoding accuracy maps were used as input for the multivariate putative NCC analysis (see below). To perform stimulus versus baseline decoding, we subsampled the stimuli trials to a 2:1 ratio with respect to baseline. The SVM cost function was weighted by the number of trials from each class. Plots were generated using Matplotlib (3.3.2)[109].

Decoding schemes for the different predictions

To test GNWT and IIT decoding predictions, stimulus category (faces versus objects and letters versus false fonts) was decoded separately for the task-relevant and task-irrelevant conditions (within-task category decoding), whereas orientation (front view versus left view versus right view) was decoded on the combined data from the two task conditions. In addition, cross-task category decoding from the task-relevant to task-irrelevant condition and vice versa was performed to test generalization by training classifiers on one condition and testing on the other condition. Both within-task category and orientation decoding were performed in a leave-one-run-out cross-validation scheme for fMRI and in an *k*-fold cross-validation scheme for MEG and iEEG.

For category decoding, trials from each task condition (that is, task relevant or task irrelevant) were extracted for each category comparison of interest: 160 face/160 objects classification, 160 letters/160 false-fonts classification within each task-relevant condition for MEG, and half the trials for iEEG. For fMRI, there were 64 trials for each category in each task-relevant condition. For orientation decoding, task-relevant and task-irrelevant trials were collapsed within category to increase the signal-to-noise ratio, resulting in 160 front, 80 left and 80 right trials per category for MEG, and half these numbers for iEEG. For fMRI, there were 64 front and 32 left and right trials per category. Decoding was evaluated using accuracy measures, tested against 50% chance level for category decoding (binary classification) and against 33% chance level for orientation decoding (three-class classification). For orientation decoding, balanced accuracy was used due to the unbalanced number of trials for the different orientations. The SVM cost function was weighted by the number of trials per class to reduce bias to the class with the highest number.

Balancedaccuracy=13(Sensitivityfront+Sensitivityright+Sensitivityleft)

For within-task decoding (for example, classification of categories across time), a classifier at each time point was trained and tested separately using a fivefold cross-validation (with three separate repeats of cross-validation). For cross-task decoding (task relevant → irrelevant and task irrelevant → relevant), each SVM model was trained on one task (for example, faces–objects in the task-relevant condition) and tested on the second task (for example, faces–objects in the task-irrelevant condition). As cross-decoding in iEEG data is performed across all pooled electrodes, an additional cross-validation step was performed on this modality data to provide a confidence metric (for example, confidence intervals) using a fivefold cross-validation with three repetitions (for example, train on 80% of task 1, and test on held-out 20% of task 2).

Within-task temporal generalization was performed by training a classifier at each time point (using selectKbest feature selection) and testing its performance across all time points using the same set of selected features and three repetitions of fivefold cross-validation. To generalize from one task to another across all time points, cross-temporal generalization was used: a classifier was trained at each time point in task 1 (for example, task relevant) using selectKbest feature selection, and tested across all time points in task 2 (for example, task irrelevant) using the same set of selected features. Cross-validation was performed in the same manner as in cross-decoding.

Additional decoding analyses were performed on all trials aligned to the stimulus onset (for example, −0.2 to 2 s relative to stimulus onset) and stimulus offset (−0.5 to 0.5 s around stimulus offset). For the latter analysis, all trials from different durations were aligned to the stimulus offset.

To assess the prediction of IIT that included prefrontal regions along with posterior regions to the decoding of categories will not significantly affect decoding accuracy, we performed an additional decoding analysis in which the decoding performance of electrodes from the IIT region were compared with the decoding performance when electrodes from both the posterior + PFC ROIs are included. The PFC ROIs included all PFC ROIs except for inferior frontal sulcus, as it belongs to the IIT extended ROIs. Posterior ROI included all IIT ROIs shown in Supplementary Table 26. The analysis compared the decoding accuracy for a model including all electrodes from posterior regions to a separate model in which electrodes (features) from posterior and PFC regions were combined (for example, feature combination). Training and testing of the individual models followed all previously described cross-validation procedures, and model comparison was performed using a variance-corrected paired *t*-test[110] and complemented with Bayesian analysis.

We also tested this prediction on the fMRI data. To select features to be used for both analyses, the face versus object contrast for each participant was masked by a predefined anatomical posterior ROIs as well as PFC anatomical ROIs, defined the same way as described above. Within each of the two ROIs, the 150 voxels that are most selective to each of the to-be-decoded stimuli were defined as the decoding ROIs (300 voxels total) for each participant. The first analysis compared the decoding accuracies for a model that included 300 voxels from the posterior ROIs as features to another model that included 600 voxels (300 features from each ROI). In the second analysis, two separate models were constructed, calibrated and combined as described above. For the two analyses, model comparison was performed using a group-level one-sample permutation test to determine if accuracies obtained by combining posterior and PFC ROIs were significantly higher than the accuracies obtained based on posterior ROIs only. FDR was used to correct for multiple comparisons. Bayesian analysis was performed to quantify evidence for the null hypothesis that adding prefrontal ROIs will not improve decoding accuracy.

Duration analysis

Neural responses were extracted from three windows of interest (0.8–1.0 s, 1.3–1.5 s and 1.8–2.0 s) and compared using LMMs. Four theory agnostic models were fitted: a null model, a duration model (three durations), a windows of interest model, and a duration and windows of interest model. Two theory models were fitted: the GNWT model predicts activation (ignition) following stimulus offset (0.3–0.5 s) independent of duration, with virtually no response in between. The IIT model predicts sustained activation for the duration of the stimulus returning to baseline after stimulus offset. Both theoretical models were complemented with an interaction term between category (faces, objects, letters and false fonts) and the theories' predictors, to account for regions showing selective responses to categories. BIC was used to define the winning model and we computed Bayes factors based on the difference in BIC values, comparing the GNWT model (with or without interaction) against either the null model (intercept only) or the time-window model (capturing amplitude changes over time)[111].

Models for iEEG were fitted per electrode on the predefined ROIs, using the high-gamma (AUC), alpha (8–13 Hz, obtained through Morlet wavelets, $f = 8$–13 Hz, in 1-Hz steps; $f/2$ cycles, AUC),and ERPs (peak to peak) as signal, separately for task-relevant and task-irrelevant condition.

MEG models were fitted to source data on the predefined ROIs, using the gamma (60–90 Hz) and alpha (8–13 Hz) bands as signal, separately for task-relevant and task-irrelevant conditions. Time-frequency analyses were performed on source-data using Morlet wavelets ($f = 8$–13 Hz, in 1-Hz steps; $f/2$ cycles; $f = 60$–90 Hz, in 2-Hz steps, $f/4$ cycles) and were baseline corrected. Spectral activity was computed for each vertex, baseline corrected and then averaged across trials within each parcel included in the ROIs, yielding a unique time course per ROI parcel. In addition, a single-source time course capturing the entire prefrontal ROI and the posterior ROI was computed by averaging the spectral activity within an ROI. Models were fitted on each parcel and ROI, as defined by the theories.

Representational similarity analysis

To examine how the neural representations evolved over time in response to the different stimulus properties (that is, category, orientation and identity representation), we performed cross-temporal RSA on source-level MEG data and iEEG high-gamma power within each of the theory-defined ROIs, using all trials. Specifically, at each set of data points, we computed a representational dissimilarity matrix (RDM) by calculating the correlation distance (1 − Pearson's r, Fisher corrected) between all pairs of stimuli (the preregistration document described a different method that was however updated to optimize trial numbers; see section 14 in Supplementary Information for justification). Next, to quantify the representational space occupied by one class versus another, we computed the average within-class distances versus the average between-class distances. This analysis was performed in a cross-temporal manner, in which RDMs were computed between all stimuli at time point t_1 and the corresponding set of stimuli at time points $t_1, t_2, \ldots t_n$.

Long trials (1.5 s) were used to investigate category and orientation representation. As specific identities were repeated a limited number of times per duration, both intermediate (1.0 s) and long (1.5 s) trials were combined and equated in duration by cropping the 1–1.5-s time interval for long trials. This was done to allow for the analysis of at least three (3) presentations of the same identity.

To evaluate the theoretical predictions about when significant content representation should occur, we subsampled the observed cross-temporal representational matrices in four time windows (0.3–0.5 s, 0.8–1.0 s, 1.3–1.5 s and 1.8–2.0 s). The subsampled

matrices were correlated to the model matrices predicted by GNWT and IIT (see Fig. 1a, right panel) using Kendall’s tau correlation. If the correlation was significant (see below) for at least one of the predicted matrices, we computed the difference between the transformed correlation ((r+1)/2) to each theory, and compared this difference against a random distribution to obtain a *P* value. If the correlation with the theory-predicted pattern in the theory ROI was significantly higher than the other model, we considered the theory prediction to be fulfilled.
To generate a null distribution of cross-temporal RSA surrogate matrices, we repeated the procedure outlined above 1,024 times, randomly shuffling the labels. Next, the observed RSA matrix was *z*-scored using the null distribution as:

$$z_{i,j}=obs_{i,j}-\mu surr_{i,j}\sigma surr_{i,j}$$

Where $obs_{i,j}$ is the observed within-versus-between class difference at time points *i* and *j*, and $\mu surr_{i,j}$ and $\sigma surr_{i,j}$ are the mean and standard deviation of the surrogate representational similarity matrix at time points *i* and *j*, respectively. Cluster-based permutation tests[112], *z*-score threshold of $z = 1.5$ for clustering, were used to evaluate significance. RSA surrogates were also used to assess the significance of the correlation between the observed matrices and the predicted matrices of the theories. First, a null distribution of possible correlations was generated for each of the theories by correlating each of the surrogate matrices to each of the theory-predicted matrices. Next, a *P* value was obtained for each theory-predicted matrix, by locating its observed correlation within the null correlation distribution. The same procedure was used to assess the significance of the difference in correlation to IIT and GNWT matrices (for example, each of the surrogate matrices was correlated to each of the theory-predicted matrices and the difference between the two was computed). *P* values were FDR corrected ($q \leq 0.05$)[75].
For iEEG, the high-gamma power per electrode within the predefined anatomical ROI was averaged in 0.02-s non-overlapping windows. Electrodes were used as features for the RDM. The data were vectorized across all electrodes within a ROI (for example, samples × significant electrodes) to compute the RDMs. A total of 576 and 583 electrodes entered this analysis for the prefrontal and posterior ROI, respectively. The resultant RDM was subjected to a PCA, and the first two dimensions were plotted against each other to produce a 2D projection of dissimilarity scores across all pairs for each of the 100 subsampling repetitions. The PCA components were aligned across repetitions using Procrustes alignment and averaged together for visualization purposes[113,114].

For MEG, the same analysis was run on the source reconstructed data within the predefined anatomical ROIs used for the decoding analysis, bandpass filtered (1–40 Hz) and downsampled (100 Hz). For the category and orientation analysis, pseudotrials and temporal moving-average methods were used to optimize the RSA analysis and improve the signal-to-noise ratio. For identity, single trials were used. Vertices within the ROIs were used as features. The statistical testing differed from that conducted on the iEEG data, as it was performed at the participant level. Similarly to the iEEG analysis, we first tested whether the correlation between the data and the model predicted by each theory was greater than zero using the Kendall's tau measure, and then compared between the theories using the Mann–Whitney U rank test on two independent samples.

Functional connectivity analysis

For both iEEG and MEG, PPC[46] was computed between each category-selective time series (face selective and object selective) and either the V1/V2 or the PFC time series.

For iEEG, the PPC analysis included electrodes in V1/V2 visual areas, in PFC ROIs (see Supplementary Table 26), and face-selective and object-selective electrodes (see 'Identification of task-responsive channels'), as long as they were 'active' during the task. As both theories predict different types of activation (for example, ignition versus sustained activation), channels were categorized as active if they showed an increase in high-gamma power relative to baseline (−0.5 to −0.3 s, $P < 0.05$, signed-rank test) evaluated across all trials (task relevant + irrelevant, intermediate + long trials, combined across both categories), for the 0.3–0.5-s window (GNWT), or in all time windows: 0.3–0.5 s, 0.5–0.8 s and 1.3–1.5 s (IIT).

For MEG, the category-selective single-trial time courses used to define the ROIs for PPC analysis were extracted using the generalized eigenvalue decomposition (GED) method[115]. Two GED spatial filters were built by contrasting either faces or objects against all other categories during the first 0.5 s after stimulus onset. Single-trial covariance matrices were computed separately for signal and reference for all vertices within the fusiform ROI identified from the FreeSurfer parcellation using the Desikan atlas[116], and the Euclidean distance between them was z-scored. Trials exceeding 3 z-scores were excluded. The reference covariance matrix was regularized to reduce overfitting and increase numerical stability. The GED was then performed on the two covariance matrices, resulting in n (=rank of the data) pairs of eigenvectors and eigenvalues. The eigenvector associated with the highest eigenvalue was selected as a GED spatial filter, which in turn was applied to the data to compute the single-trial GED component time series. A GED spatial filter was extracted also for the PFC ROI, on parcels from the Destrieux atlas[73], to identify the distributed pattern of sources that are responsive to visually presented stimuli. Specifically, a spatial filter was built by contrasting source-level frontal slow-frequency activity (30-Hz low-pass filter) after stimulus onset (0–0.5 s) against baseline (−0.5 to 0 s). V1/V2 areas were identified using the Wang Atlas[74] and a singular values-decomposition approach. For the GED, the 1.0-s and 1.5-s duration trials were used to minimize overlap with the transient evoked at stimulus onset.

PPC was computed for each MEG time series–iEEG electrode pairing, for all face trials and object trials separately. Analyses were performed on 1.0-s and 1.5-s duration trials, separately on task-relevant and task-irrelevant trials and also combined to maximize statistical power. To compute synchrony, time-frequency analysis of the broadband MEG and LFP signal was performed using Morlet wavelets (f = 2–30 Hz, in 1-Hz steps; 4 cycles; f = 30–180 Hz for iEEG or f = 30–100 Hz for MEG, in 2-Hz steps, f/4 cycles), and PPC was then computed by taking the difference in phase angle between MEG time series–iEEG electrode at each time t and frequency f for a specific trial and computing PPC across all trials in a category (for example, faces) as:

$$PPC(f,t)=\frac{2}{N(N-1)}\sum_{j=1}^{N-1}\sum_{k=j+1}^{N}\cos(\theta_j(f,t)-\theta_k(f,t)),\quad j=\{1,\ldots,N_{trials}\}$$

$\theta_{j,k}(f,t)=\theta(f,t)_{e1\,or\,GEDfilter}-\theta(f,t)_{e2\,or\,GEDfilter}$, for all frequencies f and at all times t.

For iEEG, PPC for each category-selective site was then averaged across all its pairings (for example, all PFC electrodes pairings or all V1/V2 pairings within that patient). The variability in electrode coverage across patients precluded a within-participants analysis. Therefore, to achieve sufficient statistical power, we pooled all derived PPC values from one electrode pairing (for example, face selective to the PFC) across all patients into one ROI-specific analysis. A similar approach was used on the MEG parcels.

To quantify content-specific synchrony enhancement, the difference in PPC was computed between within-category and across-category trials (for example, for face-selective sites, the change in PPC was computed between faces versus objects trials) using a cluster-based permutation test[106]. This was done for both modalities.

As an exploratory analysis, we also investigated dynamic functional connectivity using the Gaussian copula mutual information[117] approach to evaluate the dependencies between time series. This power-based measure of connectivity was implemented using the conn_dfc method from the Frites Python package[118]. We used the same parameters as for the PPC analysis, with the following exceptions: for both MEG and iEEG, power was estimated through a multitaper-based method (using a frequency-dependent dynamic sliding window: 2–30 Hz, T = 4 cycles; 30–100 Hz, $T4/f$ using a 0.25-s sliding window). For iEEG, the high-frequency range was extended from 30 to 180 Hz, $T = 4/f$ cycles). DFC was performed per frequency band, 0.1-s sliding window and 0.02-s steps.

For fMRI, connectivity was assessed through gPPI implemented in SPM[119]. The FFA and lateral occipital cortex were defined as seed regions per participant based on an anatomically constrained functional contrast. Anatomically, FFA seeds were constrained to the 'inferior occipital gyrus (O3) and sulcus' and 'lateral occipito-temporal gyrus (fusiform gyrus; O4–T4)'. LOC seeds were constrained to the 'middle occipital gyrus (O2; lateral occipital gyrus)' and the 'middle occipital sulcus and lunatus sulcus' (Destrieux ROIs 2 and 21 for FFA, and ROIs 19 and 57 for LOC; see 'Anatomical ROIs').

Candidate seed voxels within the above-mentioned anatomical ROIs were defined as those with $z > 1$ in the contrast of parameter estimates of all stimuli versus baseline. Three participants with less than 300 candidate seed voxels were excluded from the analysis. This was done to ensure that the seed voxels were visually driven. Next, using an unthresholded contrast of parameter estimates between 'relevant and irrelevant faces' and 'relevant and irrelevant objects', the 300 voxels most responsive to faces within the FFA anatomical ROIs were selected for the FFA seed, and the 300 voxels most responsive to objects within the LOC anatomical ROIs were selected for the LOC seed.

gPPI analysis was performed per participant and seed region separately, including an interaction term between the seed time-series regressor (physiological term) and the task regressor (psychological term) at the participant-level GLM[119], separately for task-relevant and irrelevant conditions, and also combining across tasks to increase statistical power. For combined conditions, the model design matrix for each participant included regressors for task-relevant and task-irrelevant faces, objects, letters and false fonts collapsed across conditions (four regressors) as well as a regressor for targets (irrespective of their category), yielding five regressors in total. As for separated conditions, the model design matrix included regressors for task-relevant and task-irrelevant faces, objects, letters and false fonts (eight regressors) as well as a regressor for targets (irrespective of their category), yielding nine regressors in total. For each seed, group-level analysis was performed using a cluster-based permutation test (preferred over the preregistered FDR correction), complemented by Bayesian analysis. See section 14 in Supplementary Information for a justification of this change to evaluate the statistical significance of face > object contrast parameter estimates across participants ($P < 0.05$).

Putative NCC analyses

A series of conjunction analyses were performed on the fMRI data to identify (1) areas responsive to task goal, (2) areas responsive to task relevance, and (3) areas putatively involved in the neural correlates of consciousness. We note that the contrasts proposed below might overestimate the neural correlates of consciousness and that the fast-event-related design adopted here might be suboptimal to detect activity changes in the salience network[120], that is, potentially underestimating some regions that might be involved in conscious processing. We therefore have adopted a conservative approach that distinguishes between areas that might participate in consciousness versus those that definitely do not.

The conjunction defining areas responsive to task goals was defined as [TaskRelTar > bsl] and [(TaskRelNonTar = bsl) and (TaskIrrel = bsl)]. This contrast captures areas that show an increase of BOLD signal for targets but not for other stimuli. The following conjunction identified areas responsive to task relevance: [(TaskRelTar > bsl) and (TaskRelNonTar ≠ bsl)] and [TaskIrrel = bsl]. This contrast identifies areas displaying differential activity for all task-relevant stimuli, but are insensitive to non-task-relevant stimuli. Finally, the following conjunction was used to identify the putative NCC areas: [(TaskRelNonTar (stim id) > bsl) and (TaskIrrel (stim id) > bsl)] or [(TaskRelNonTar (stim id) < bsl) and (TaskIrrel (stim id) < bsl)], critically detecting areas that are responsive to any stimulus category irrespective of task, with consistent activation or deactivation. Thus, this analysis casts a wide net to identify areas that can potentially be the neural correlates of consciousness, whereas excluding areas that do not respond to task-relevant or irrelevant stimuli (meaning that areas that respond both to the task and to the content of perception are still included).

To compute conjunctions, we first ran a GLM (see above) corrected for multiple comparisons (Gaussian random-field cluster-based inference). Equivalence to baseline was established using a JZS Bayes factor test, with a Cauchy prior (r scale value of 0.707, as implemented in Pingouin (0.5.1)[121]. Evidence maps were thresholded at $BF_{01} > 3$. The thresholded z maps and the Bayesian evidence maps on the group level were used for the conjunction analysis. For conjunctions including an ‘unequal to’, a ‘logical and’ operation was used between the directional z maps, after thresholded maps were binarized. For the putative NCC contrast, conjunctions were performed separately for activations and deactivations, using a ‘logical and’ operator for the task-relevant and irrelevant z maps. The resulting maps were combined using a ‘logical or’ operation to discard areas showing effects of opposite direction for task-relevant and task-irrelevant stimuli. This analysis was also done at the participant level, masked using the anatomical ROIs, to account for inter-participant variability. For each ROI, the proportion of participants with voxels included in the conjunction is reported. The multivariate version of the putative NCC analysis was done using the thresholded statistical maps obtained from the whole-brain searchlight decoding based on a participant-level stimulus versus baseline-decoding accuracy maps (for details regarding the decoding approach used, see ‘Decoding analysis’).

Scientists explore where consciousness arises in the brain

FILE PHOTO: People are silhouetted against the setting sun on top of the Drachenberg in Berlin, Germany, Germany, August 19, 2019. REUTERS/Fabrizio Bensch/File Photo© Thomson Reuters

By Will Dunham

WASHINGTON (Reuters) -Consciousness is at the center of human existence, the ability to see, hear, dream, imagine, feel pain or pleasure, dread, love and more. But where precisely does this reside in the brain? That is a question that has long confounded scientists and clinicians. A new study is offering fresh insight.

In a quest to identify the parts of the brain underpinning consciousness, neuroscientists measured electrical and magnetic activity as well as blood flow in the brains of 256 people in 12 laboratories across the United States, Europe and China, while the participants viewed various images. The measurements tracked activation in various parts of the brain.

The researchers found that consciousness may not arise in the "smart" part of the brain - the frontal areas where thinking is housed, which progressively grew in the process of human evolution - but rather in the sensory zones at the back of the brain that process sight and sound.

"Why is any of this important?" asked neuroscientist Christof Koch of the Allen Institute in Seattle, one of the leaders of the study published this week in the journal Nature.

"If we want to understand the substrate of consciousness, who has it - adults, pre-linguistic children, a second trimester fetus, a dog, a mouse, a squid, a raven, a fly - we need to identify the underlying mechanisms in the brain, both for conceptual reasons as well as for clinical ones," Koch said.

The subjects in the study were shown images of people's faces and various objects.

"Consciousness is the way it feels like to see a drawing of a toaster or Jill's face. Consciousness is not the same as the behavior associated with this feeling, for example pushing a button or saying, 'I see Jill,'" Koch said.

The researchers tested two leading scientific theories about consciousness.

Under the Global Neuronal Workspace Theory, consciousness materializes in the front of the brain, with important pieces of information then broadcast widely throughout the brain. Under the Integrated Information Theory, consciousness emanates from the interaction and cooperation of various parts of the brain as they work collectively to integrate information that is consciously experienced.

The findings did not square with either theory.

"Where are the neuronal footprints of consciousness in the brain? Very crudely put, are they in the front of the cortex - the outermost layer of the brain - such as the prefrontal cortex, as predicted by the Global Neuronal Workspace Theory?" Koch asked.

It is this prefrontal cortex that makes our species uniquely human, driving higher-order cognitive processes such as planning, decision-making, reasoning, personality expression, and moderating social behavior.

"Or are the footprints in the back regions of the cortex, the posterior cortex?" Koch asked. The posterior cortex houses the regions where hearing and vision processing occurs.

"Here, the evidence is decidedly in favor of the posterior cortex. Either information pertaining to the conscious experience couldn't be found in the front or it was far weaker than in the back. This supports the idea that while the frontal lobes are critical to intelligence, judgment, reasoning, etc., they are not critically involved in seeing, in conscious visual perception," Koch said.

However, the study did not identify enough connections that last for as long as the conscious experience in the back of the brain to uphold the Integrated Information Theory.

There are practical applications in gaining a deeper understanding of the mechanics of consciousness in the brain.

Koch said it would be important for how doctors deal with patients in a coma or patients in a vegetative state or with unresponsive wakefulness syndrome, when they are awake but present no signs of awareness due to traumatic brain injury, stroke, cardiac arrest, a drug overdose or other causes.

"If the patient remains in this unresponsive state for longer than a few days without signs of recovery, the clinical team initiates discussion with the family around, 'Is this what they would have wanted?'" Koch said.

Of such patients, 70% to 90% die because a decision has been made to withdraw life-sustaining treatment.

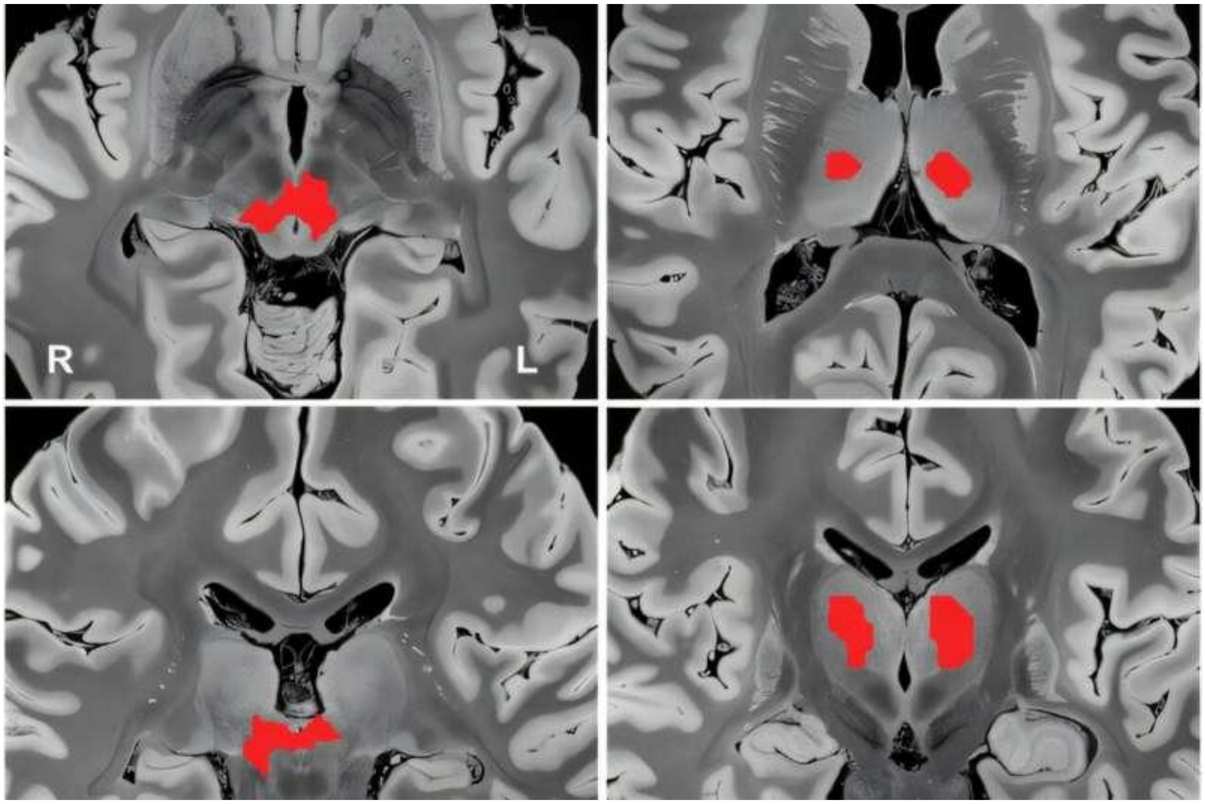
"However, we now know that around a quarter of patients in either coma or vegetative state/unresponsive wakefulness syndrome are conscious - covert consciousness - yet are unable to signal this at the bedside," Koch said, referring to research published last year in the New England Journal of Medicine. "Knowing about the footprints of consciousness in the brain will let us better detect this covert form of 'being there' without being able to signal."

(Reporting by Will Dunham, Editing by Rosalba O'Brien)

MAY 15, 2025

Deep brain regions link all senses to consciousness, study finds

by Karen Guzman, Yale University



Midbrain and central thalamus show shared subcortical early activations (increases), observed in 11 tasks across four sensory modalities, including, vision, audition, taste, and touch. Credit: *NeuroImage* (2025). DOI: 10.1016/j.neuroimage.2025.121224

A Yale-led study shows that the senses stimulate a region of the brain that controls consciousness—a finding that might inform treatment for disorders related to attention, arousal, and more.

Humans perceive and navigate the world around us with the help of our five senses: sight, hearing, touch, taste and smell. And while scientists have long known that these different senses activate different parts of the brain, a new Yale-led study indicates that multiple senses all stimulate a critical region deep in the brain that controls consciousness.

The study, published in the journal *NeuroImage*, sheds new light on how sensory perception works in the brain and may fuel the development of therapies to treat disorders involving attention, arousal, and consciousness.

In the study, a research team led by Yale's Aya Khalaf focused on the workings of subcortical arousal systems, brain structure networks that play a crucial role in regulating sleep-wake states. Previous studies on patients with disorders of consciousness—such as coma or epilepsy—have confirmed the influence of these systems on states of consciousness.

But prior research has been largely limited to tracking individual senses. For the new study, researchers asked if stimuli from multiple senses share the same subcortical arousal networks. They also looked at how shifts in a subject's attention might affect these networks.

For the study, researchers analyzed fMRI (functional magnetic resonance imaging) datasets collected from 1,561 healthy adult participants as they performed 11 different tasks using four senses: vision, audition, taste, and touch.

They made two important discoveries: that sensory input does make use of shared subcortical systems and, more surprisingly, that all input—regardless of which sense delivered the signal—stimulates activity in two deep brain regions, the midbrain reticular formation and the central thalamus, when a subject is sharply focused on the senses.

The key to stimulating the critical central brain regions, they found, were the sudden shifts in attention demanded by the tasks.

"We were expecting to find activity on shared networks, but when we saw all the senses light up the same central brain regions while a test subject was focusing, it was really astonishing," said Khalaf, a postdoctoral associate in neurology at Yale School of Medicine and lead author of the study.

The discovery highlighted how key these central brain regions are in regulating not only disorders of consciousness, but also conditions that impact attention and focus, such as attention deficit hyperactivity disorder. This finding could lead to better targeted medications and brain stimulation techniques for patients.

"This has also given us insights into how things work normally in the brain," said senior author Hal Blumenfeld, the Mark Loughridge and Michele Williams Professor of Neurology who is also a professor in neuroscience and neurosurgery and director of the Yale Clinical Neuroscience Imaging Center. "It's really a step forward in our understanding of awareness and consciousness."

Looking across senses, this is the first time researchers have seen a result like this, said Khalaf, who is also part of Blumenfeld's lab.

"It tells us how important this brain region is and what it could mean in efforts to restore consciousness," she said.

More information: Aya Khalaf et al, Shared subcortical arousal systems across sensory modalities during transient modulation of attention, *NeuroImage* (2025). DOI: 10.1016/j.neuroimage.2025.121224
Journal information: NeuroImage
Provided by Yale University

What (If Anything) Physics Can Say About Consciousness Denyse O'Leary June 13, 2025, 6:44 AM

Photo credit: Basile Morin, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons.

Theoretical astrophysicist Ethan Siegel has a question: Does physics truly have anything to say about consciousness? At *Big Think,* he offers some observations from a materialist perspective that are worth commenting on.

He starts with a number of questions and then makes two declarations:

The Questions:

What does it truly mean to be conscious? Where does consciousness come from? Are humans the only conscious species, or do other animals, non-animal forms of life, or even non-living things possess some form of consciousness? While many have opined and put forth hypotheses on the matter, it remains a mystery.

"DOES PHYSICS TRULY HAVE ANYTHING TO SAY ABOUT CONSCIOUSNESS?," MAY 14, 2025

Consciousness is very hard to define or pin down when approached in the manner of the first two questions. It's not so hard when approached experientially, as it is in the third question.

Many animal species show consciousness in philosopher Thomas Nagel's sense: There is something that it "is like" to be a dog or a cat. That's *sentient* consciousness — awareness of circumstances and feelings. But perhaps, by contrast, there is nothing that it "is like" to be a sand dollar or a sea cucumber. The fact that some animals have sentient consciousness does not show that all do.

We also know by experience that human consciousness includes the capacity, when fully developed, for reason and moral choice, which is a higher order of consciousness than sentience.

What about objects? If we want to argue that they have consciousness too, we need to define the term differently from the experiential way we usually do.

For example, if I tell you that two dogs, Bowser and Scout, get along well, I am telling you something different from saying, perhaps, that two lamps suit the rec room decor. In the first instance, I am telling you something about the inner states of Bowser and Scout, as inferred by a human. But in the second instance, I assume that no lamp has any inner state and it is a matter of my human judgment whether they outwardly suit the decor.

Let's go back to the first two questions for a moment: What does it truly mean to be conscious? and Where does consciousness come from? The first question can't be answered generally because the answer might be different for many different conscious beings. The second is easily disposed of: We have no idea, at present, how consciousness originates. But we can say, if we dare, that materialist theories do not appear to be shedding much light.

Now for His Declarations

Here's what physics — the most fundamental of all the sciences — has to say about consciousness.

At the very core of the matter are two basic ideas:

- *the idea that we live in a material reality, and that everything that exists in our material reality can be described in terms of, well, the constituent parts of reality that exist in space and time,*
- *and the idea that any phenomenon — including consciousness — can be rigorously defined and put to experimental, observational, and/or measurable tests.*

To a physicist's way of thinking, these are non-negotiable starting points for attempting to gain a physical understanding of any phenomenon in the Universe.
"ABOUT CONSCIOUSNESS?"

But if the Ship Sinks, So Does the Cargo…

In short, Siegel seeks an account of consciousness that places it squarely within a materialist framework. The problem is that, as Michael Egnor and I show in *The Immortal Mind* (Worthy, June 3, 2025) — and others have shown in similar works — consciousness is not a material thing and cannot be put into a materialist framework. Only a distorted understanding of consciousness would emerge.

But, in fact, a number of thinkers have tried to do precisely that: Their slogan is typically some variant of "The mind is simply what the brain does." The overall point of view is called eliminative materialism.

The trouble is, eliminative materialism is *itself* a project of consciousness. And if the ship sinks, so does the cargo.

Siegel appears not to understand the problem clearly. He closes with:

Consciousness is a very difficult puzzle: one that is difficult to even define, much less to solve. But it is just as much a part of our physical reality as anything else we interact with, and any approach that asserts otherwise has a fatal flaw from the outset: it's already abandoned science.
"ABOUT CONSCIOUSNESS?"

Let’s stop and take this in. Science is, in the first instance, a product of human consciousness. It gains its prestige from the fact that its correct predictions and the reasoning behind them form an *immaterial intellectual construct* in human minds. The scientist implicitly understands that her reasoning in these matters is something that stands apart from the material things she is reasoning about. Otherwise, it would not be reasoning at all; it would just be more data.

In that sense, her consciousness is not “just as much a part of our physical reality as anything else we interact with.” As philosopher Edward Feser has warned, eliminative materialism is a snake that eats its own tail. There is no escape, on its own terms, from its dilemma.

Abandoning Science?

If there is talk of abandoning science, the first thing that must be said is this: If science, as interpreted by some, cannot address immaterial realities like our ability to reason and thus *do*science — on their own terms — then so much the worse for that interpretation of science.

Many people go into science (or other fields) hoping that it will validate what they believe about reality. Some are rewarded and others are not. But validating one’s view of reality is not the purpose of science. Insisting that it do so tends to limit its growth.

The Art of Shifting Your Reality:
4 Deep Practices for Clearing Emotional Blockages and Harnessing the Power of Your Subconscious Mind

BY JOSEPH DRUMHELLER

The keys to altering your reality lie deep within the subconscious mind.

Joy, happiness, excitement, passion, purpose, meaning, peace, contentment, fulfillment, abundance, freedom, etc. How many of us get to experience these words, let alone live them out on a daily basis? Not nearly enough.

The adventure we're about to embark on will lead into the realm of these words. It's the sole reason for taking up a spiritual path: to increase the quality of life. Try to keep that in mind.

A friend of mine once told me, "There are two kinds of people in this world, those who say 'I can' and those who say 'I can't.' And ya know what? They're both right." The events in our lives are outward projections of what's lurking in our subconscious mind. That includes our thoughts, feelings, perceptions, and beliefs. We are who we think, feel, and believe we are. That goes for the good and the bad. If that's so, can we intentionally put thoughts, feelings, perceptions, and beliefs into the subconscious mind to create our future? The answer is: You bet! It's called manifesting.

"Whoever you think you are as person and however you think the world operates will be reflected back to you in the events of your life."

Manifesting is simply utilizing the creative force of the subconscious mind by directing into the future. Anything is possible! The only limitation is what you can imagine. It's like applying a powerful subconscious healing technique to the future. It can be applied to almost anything: healing, sports, technical skills, physical abilities, relationships, finances, health, etc. Roadblocks to manifesting show up as resistance and/or subconscious emotional charges. You'll need to use subconscious healing and reprogramming techniques outlined later in this article to get rid of those.

And that's why I'm here. I'm here to guide you through the process and help you avoid bumps along the way. The effort you put in will produce noticeable results in your daily life. It will also influence the lives of everyone you come in contact with. So if you're interested in making your life and the world a better place, then by all means, please join me.

The Conscious Mind—Noise, Noise, Noise!

Each and every one of us on planet Earth walks around in this little bubble called the human mind. On the outer layer of the bubble is the conscious mind, a very important and influential part of our existence that differentiates us from animals. It's our "awake" state of mind, the state where we spend the majority of our waking hours. It's the home of our cognitive and rational thinking process and our ability to organize and categorize. Much of our society is steeped in the conscious mind, including science and modern medicine.

The following list is borrowed from Yvonne Oswald's book, *Every Word Has Power*, detailing features of the conscious mind. The conscious mind:

1. Is aware of what it perceives.

2. Contacts reality through the five senses.

3. Gathers and sorts information.

4. Communicates to the Divine through the subconscious.

5. Thinks deductively.

6. Makes choices and judgments.

7. Reviews information and draws conclusions.

8. Makes generalizations.

9. Likes to analyze and categorize.

10. Requests information from the subconscious.

The conscious mind is an incredible thinking machine. It just doesn't stop. My favorite description of the mind's activity is written by Eckhart Tolle in *A New Earth*: "Thinking isn't something you do; thinking is something that happens to you." Many of these thoughts are unfocused noise and chatter. Mind chatter can range from mildly annoying to downright debilitating. There are numerous self-help books out there that encourage us to train ourselves to think positively. However, due to the sheer volume of thought we have, our control over the human mind and our mental chatter is fairly limited.

Simply recognizing mental chatter and the effect it has on your life is the first step in learning to dissipate its power. If the chatter is excessively noisy, there is likely a strong underlying emotional charge generating it from the subconscious mind.

The Subconscious Mind

Underlying the conscious mind in the human mind bubble is the subconscious mind. By definition, we are not normally aware of this aspect of our being (sub means under). While we are busy going about our day, deeply engrossed in the conscious mind, the subconscious mind lies quietly beneath the surface. It's the home of the nonphysical realms of the human experience. Some aspects include our dreams, emotions, ideas, and imagination.

The following list is borrowed from Yvonne Oswald's book, *Every Word Has Power*, detailing features of the subconscious. The subconscious mind:

1. Operates the physical body.

2. Has a direct connection with the Divine.

3. Remembers everything.

4. Stores emotions in the physical body.

5. Maintains genealogical instincts.

6. Creates and maintains least effort (repeating patterns).

7. Uses metaphor, imagery, and symbols.

8. Takes direction from the conscious mind.

9. Accepts information literally and personally.

10. Does not process negative commands.

In addition to those features listed above, the subconscious mind also holds our perceptions, subconscious beliefs of who we think we are as people and how we think the world operates. In other words, whoever we think we are as people and however we think the world operates quietly lives under the conscious mind in the subconscious. Our perceptions may be wonderful or painful. The subconscious mind does not evaluate or judge these perceptions; it simply accepts them as truth. Furthermore, the job of the subconscious mind is to play these perceptions out in the circumstances of our lives. Creating our perceptions is what the subconscious mind is designed to do. Whoever you think you are as person and however you think the world operates will be reflected back to you in the events of your life.

Whoa…wait a minute! The first time I heard that, two things went off in my head: one good and one not so good. First off, it made sense. Second, it explained why my life resembled a train wreck. I felt doomed from the get-go, helpless to do anything about it. If you feel the same way, hang on a minute. Don't panic…yet. Keep reading. It gets better.

The Critical Faculty of the Mind

Because the subconscious mind is such a continuously creative machine and the conscious mind is limited in its scope, this puts us all in a very vulnerable situation. If every subconscious perception we experienced were to be acted out in our daily lives, we would be living nothing but the Chaos Theory in motion! Just like traffic in India! Therefore, somewhere in the wisdom of creating the human experience, the critical faculty of the human mind was established to prevent us all from going absolutely nuts.

The critical faculty of the human mind is an invisible, protective barrier living somewhere between the conscious and subconscious minds. Its function is to constantly evaluate perceptions that are projected in our direction: in thought, word, or action. As perceptions are aimed our way, the critical faculty will judge it as yes or no.

Yes means the projection is in harmony with what lives in the subconscious mind. No means the projected perception is not in harmony with what's in our subconscious mind. When a yes perception arrives, the critical faculty will open up and allow the perception in, thereby allowing the existing perception to grow. When a no perception shows up, the critical faculty will remain up, rejecting the perception as an untruth.

“The incredibly strange part of the critical faculty is that it appears to be fully established by the time we’re about four years old.”

Let me give you an example: If I teach a workshop and someone comes up to me and says, “Wow that was great. You’re a wonderful teacher,” my critical faculty will let down and permit that perception to enter. My subconscious perception of I’m a wonderful teacher (i.e. I’m a wonderful person) will grow and continue to manifest in the events of my life. However, if someone in my class is sitting in the back row thinking, Boy, this guy’s an idiot, my critical faculty will remain up and reject that projection as an untruth. It simply does not fit my perception of myself or how I view the world.

the power of the subconscious mind over our perceptions and reality is immense.

The incredibly strange part of the critical faculty is that it appears to be fully established by the time we’re about four years old. What?!? The first time I read that, I almost broke down and cried. What a gyp! You mean, by the time we’re only out of our toddler years, whoever we think we are as people and however we think the world operates is nearly set in stone? And the shaping of us poor little human beings is totally at the mercy of our immediate families and environment? Afraid so. It’s a case of karma. For some reason, who knows why, our immediate environment is the integral foundation of our human experience. It’s where our learning begins.

Negative Perceptions: Emotional Charges

I have seen negative or painful perceptions referred to as “charged emotions” and the “pain body” by Oneness University and Eckhart Tolle, respectively. I prefer the Oneness term because it most accurately fits the healing work I have done. So from here on out, I’ll refer to painful subconscious perceptions as charged emotions or emotional charges.

Let’s take a look at Suzie. She’s a two-year-old girl, and today is a major red-letter day in her life. She gets to drink milk out of a real glass—not one of those little-kid plastic sippy cups, but a real glass container. Suzie’s mom is at the stove cooking dinner. She’s had a hectic day, working a job that barely enables her to make ends meet. She’s been a single mom for the past eighteen months and hasn’t seen a dime of child support for

Suzie or her two siblings. Needless to say, mom's life is on stress overload. As she cooks dinner, low and behold, Suzie drops her glass. Milk and shattered glass go everywhere. Mom, at her wit's end, spins toward Suzie and screams, "Damn it! You'll never amount to anything!!!" Blink. A powerful negative perception, or charged emotion, has just been created in Suzie's subconscious mind. The critical faculty has also built a wall around that perception to protect it. Does that mean Suzie will grow up to be worthless? Of course not! However, if she continues to receive the same message over and over again as she grows up (in thought, word, or action), the critical faculty will continue to let down, allowing that charged emotion to grow—to the point where it will eventually begin to appear in the events of her daily life. Perhaps she won't be able to hold down a job; she may go from one failed relationship to the next; or it may show up as a physical illness or injury. The subconscious mind is very powerful and can be very creative when it comes to expressing emotional charges, and it doesn't stop…until the charges are gone.

So you may be thinking, What, exactly, is an emotional charge? Subconscious emotional charges are un-experienced, subconsciously repressed feelings that produce pain. Examples include hurt, rage, anxiety, rejection, shame, unworthiness, fear, depression, and abandonment, and the list goes on and on. These are like living entities in our bodies, and they won't go away until they are experienced fully.

Human Suffering and the Power of the Mind

Let's take subconscious emotional charges one step further and journey into the world of suffering. First, it's essential to understand what suffering actually is. Is freedom from suffering the end of pain, anger, and unpleasant emotion, replaced by everlasting bliss? Sorry, but no. That wouldn't be very realistic or very human. Suffering is not pain, but the avoidance of it. Let me say that again: The definition of suffering is the avoidance of pain. Pain demands attention, and when it is avoided, it repeats itself over and over again. Suffering is the experience of the same repeating pain. Herein resides the grand paradox. In order to overcome suffering, you must experience your pain fully.

Human beings are designed for growth and experiencing life to the fullest. That includes all of our emotions. Pain, jealousy, anger, enthusiasm, joy, contentment, etc., are all part of the human experience. Look at small child who drops his/her ice cream cone on the sidewall. A few moments of intense agony are guaranteed. However, in just a few minutes, after the emotion has been experienced, the tears dry, and the child is back to experiencing life in the moment. When emotions are experienced fully, they dissipate. The problem arises when we avoid experiencing unpleasant emotions. We essentially put them in the closet to reemerge at a later date…and they do.

“Human beings are designed for growth and experiencing life to the fullest. That includes all of our emotions.”

How many of us know someone who has gone from relationship to relationship, experiencing the same problems? How about the person who avoids conflict with authority to the point of becoming seriously ill? They are just avoiding unpleasant feelings. We’ve all done it. In fact, it’s a natural response. Unpleasant emotions tend to get repressed into the subconscious mind without conscious knowledge or intention on our part. When that happens, they get stuck in the subconscious. The subconscious mind then continues to recreate uncomfortable circumstances in our lives so that the repressed emotions can be experienced. Without awareness, this saga can last a lifetime.

Repeating Patterns—Props in Your One-Person Drama

I’m a fairly sensitive guy. I have emotions running through me all the time. When I first became aware of my internal process, my big question was: How do I know if I have an emotional charge or just a random feeling? The answer is simple: repeating patterns. If you have difficult emotions repeating themselves in similar circumstances, regardless of the people you are dealing with, it’s a subconscious emotional charge. Examples may include a woman from an abusive family who keeps finding an abusive partner or a man with an impoverished father that can’t ever seem to get his act together with finances.

The combinations and possibilities are endless. There’s a potential charge for every difficult life circumstance. It takes a certain commitment to honest self-reflection to discover that your repeating patterns are actually your own subconscious invention! Emotional charges are never about the other person. They are your creation and experience only. Other people who trigger your emotional charges are only props in your own one-person drama.

Tragic Events

Let’s dive straight into an atheistic argument, a question that can rack the faith of even the most devout believer. What about suffering caused by birth defects, starvation, tsunamis, landslides, earthquakes, and disease? How does this nifty little subconscious mind and suffering theory explain random cataclysmic events of tragedy? The truth is, it doesn’t. Remember, for some reason, pain is part of the human experience. However, learning to experience pain prevents it from recurring over and over again. Sometimes it’s not easy, but as we will discover, it’s the road to the Divine.

Suffering Is Not Real: The Illusion

When we begin to grasp the notion that suffering is repeating patterns caused by subconsciously repressed pain,the light bulb should go on that says, ”Hey, wait a minute… if suffering is just repeating patterns, then it’s really based in the past.” Ding-ding! We have a winner.

In a way, suffering isn’t actually real. The past doesn’t exist. Suffering is from the past. You do the math. How much of our behavior and decision-making is based on something that technically isn’t real? A lot! Every time someone reacts to a recurring difficult situation, it’s just a response to the past. It’s like fighting ghosts; there’s nothing there but a mental image and a feeling. The problem is that those images and feelings can be extremely strong.

Here’s a fairly typical scenario I’ve seen a number of young adults go through. In fact, I was guilty of it myself throughout my twenties. For entertainment’s sake, let’s call this young buck Hugo. He was a recent transplant from Florida to my little town in the Pacific Northwest. I met Hugo through my circle of paddling pals, so we went out a few times and hit the river together. Getting to know him, I discovered he came from an extremely abusive past. He hated his old man and, as far as I could figure, with good reason. That was why Hugo picked up and moved to the Northwest, “to get as far away from that son-of-a-!!!!! as possible.”

I also heard Hugo talking about his employer. If I hadn’t known the context of the conversation, I would have sworn he was talking about his father. It wasn’t too long until he moved to another town about 150 miles away, stating something to the effect of “My boss is a jerk” and “I’m outta here.” A few months later, he moved to Alaska. Last I heard, Hugo was in Australia, still fighting the same issues with authority figures (i.e. his father).

This is a case of repeating patterns caused by a powerful subconscious emotional charge created in the past. The most effective way for Hugo to break his repeating pattern of misery is to go into the subconscious mind and heal his emotional world.

Childhood — It's Not Your Fault

Most emotional charges are programmed into your subconscious mind by someone else or some outside event. When? When we're most vulnerable, of course. Emotional vulnerability is at its peak before the critical faculty of the mind is developed. As you may recall,that takes place in very early childhood, when the subconscious mind is completely exposed.

We enter this world as pure vessels, wide open and absorbing life like a sponge. This includes the good, the bad and the ugly. Hey, nobody said life was fair. The stork settles a few of our darlings into lovely, cushy feather beds.

"We enter this world as pure vessels, wide open and absorbing life like a sponge. This includes the good, the bad and the ugly."

Others get dumped into this world under absolutely horrendous circumstances. Wherever we land is where our journey begins, like it or not, and our emotional charges develop at the whim of our surroundings.

It's not a difficult stretch to come to the conclusion that I can't emphasize strongly enough: Your suffering is not your fault. You did not create your problems, dysfunction, or disease. You are only responding to circumstances at your given level of awareness.

Let's take it one step further. In regard to the person(s) who planted charges in you, it's not their fault either. They were also responding to their charges at their given level of awareness. Take it as far back as you need to go. Who is ultimately responsible? Nobody. Suffering isn't anybody's fault; it just is. You may not believe it now, but suffering is actually what paves the road to the Divine.

Ancestry — It's Truly Not Your Fault

We inherit all sorts of stuff from out family: appearance, size, shape, character, behavior, mentality, and if we're lucky, a little money. We never asked for any of it. It's just what's on the menu. We also inherit from our ancestry through the subconscious mind, including emotional charges.

We all know someone who's a spitting image of their parents, a regular clone, or a chip off the ol' block. I knew a girl in high school whose mother's maiden name was Smith.

She later married a guy named Smythe, who ironically (or maybe not so ironically) resembled her dad. Her mother had three kids, all girls. She had three kids, all boys. Listing the similarities between those two would take all day. If you take a look within your own family or your friends' families, recognizing generational parallels will begin to be surreal.

For some, inherited ancestral traits can skip a generation. I have an acquaintance whose grandfather was a bit of a bigwig powerbroker, married three times. My acquaintance never met him and resented growing up in his shadow.

Oddly enough, he's at the beginning of his third marriage and his life is moving onto a significant public platform. His level of influence was born into him. He's had to do quite a bit of soul-searching and forgiveness to accept himself as a part of his heritage. The more he has accepted himself and his pedigree, the more his destiny of success has unfolded.

"So what's the point? We inherit traits from our families. Everybody knows that." Not only do we inherit physical and behavioral characteristics, but we also subconsciously inherit emotion. And what does that mean? It means we also inherit suffering. Some of our suffering is caused by an unusual subconscious connection to our heritage.

Understanding this possibility opens the potential for subconscious healing to a whole new level. So far, we've learned that suffering is basically not our fault. Furthermore, if we inherit suffering from ancestors who were dead and buried long ago, then it's really not our fault.

Gestation

"I've always been close to my mom, but I never knew how close." Before birth we are our mother. We're connected physically, physiologically, emotionally, and subconsciously. It's not uncommon to absorb our mother's subconscious emotional charges while we're in the womb, especially if the contributing events are traumatic.

Here’s a hypothetical scenario I’ve seen repeated is a variety of situations. Since most of us have had spouses and lovers at least once or twice in our lives, I’ll use the context of an intimate relationship, something we can all relate to.

A young woman came to see me for recurring anxiety attacks that were happening at an accelerated and stronger rate. They exploded into her life whenever she crossed some ambiguous line of going from a romantic to a committed relationship. It had ruined several previous relationships.

“Before birth we are our mother. We’re connected physically, physiologically, emotionally, and subconsciously.”

She was now involved with a man she genuinely loved. The attacks were becoming so frequent and so strong that she could barely function. She was desperate. This was in the early days of my practice, back when I took more of an analytical approach to subconscious healing.

Over the course of a few sessions, we probed into the events of her childhood and had some initial satisfactory success. However, the anxiety attacks still came back, albeit with less frequency and intensity. I couldn’t put my finger on what was behind this. We just didn’t seem to be getting to the core of the problem.

In conversation, I asked about her parents, and that led to the topic of their divorce. The split happened rather abruptly, and it was rather grueling on her mother. I asked her how old she was when her parents broke up, to which she replied, “Oh, I wasn’t born yet. I was still in my mom’s tummy.” Flash! Maybe she picked up emotional charges from her mom while she was pregnant. By the next session, her anxiety attacks gone… permanently.

Family, Societal, and Media Conditioning

“All of us Kennedys are politicians… To get a good job, you need a good education…My TV would never lie to me.” Family, societal, and media conditioning is lathered upon us from the moment we enter this world.

What is conditioning? It's received input through word,thought, or action that influences our behavior. When we receive a message over and over again (especially when we're very young), it enters the subconscious mind and will eventually be acted out in our lives. It's not necessarily good or bad. Problems arise when the input received from our surroundings (family, society, media) doesn't fit who we areas people. This creates internal conflict and puts us at odds with ourselves and our world, which is — you guessed it —a powerful subconscious emotional charge.

Family

"We've been lawyers in this family as long as I can remember. My dad, his dad, his dad's dad. We go back generations. By God, you'll be one too!" That may be fine for my brother Seymour, but I'm just not wired that way. I need to beat to my own free-spirited drum. For me, becoming a lawyer would be a straight shot to hell.

Families can be blown apart by adherence to family conditioning, especially if it supersedes genuine needs. How many of us know someone who has gone into the family business and hated every minute of it? How many of us have been crushed by family disapproval over something we were genuinely passionate about? If we are to grow and mature into the people we are meant to be in this life, we need to listen to the voice within. It may or may not be in accordance with family approval. Blind adherence to family conditioning, at the expense of yourself, can produce a lifetime of suffering.

Society

"Man is mad." That's one of my favorite quotes I picked up at Oneness University. Because each of us projects the circumstances of emotional charges into our lives, many societal values are tainted with the spewing of these charges. It can be a vicious cycle that propagates even more suffering.

"In order to transform our world, we must transform ourselves on the inside. As we do, our values will shift."

Collective values get filtered through the charges and chattering minds of millions. Ideals behind a societal system can easily be polluted into the perpetuation of things like needless war and excessive corporate greed. It's no wonder the state of affairs can get so messy. As each of us goes, so goes society. In order to transform our world, we must transform ourselves on the inside. As we do, our values will shift. Emphasis will then be placed more on the quality of life and less on the trappings of the conscious and subconscious mind and unresolved emotion.

Media

Technology? Talk about a runaway freight train. Don't get me wrong, I love it—or at least parts of it. I can't denounce technology as I sit here pecking away on my laptop. Let's admit it: we're hooked. What's worse is that our kids are hooked. Most kids nowadays are more familiar with some flat, illuminated screen than members of their own extended families. If I were to ask my sixteen-year-old stepdaughter today, "What was the weather like this morning?" or "What phase is the moon in right now?" she would probably answer, "Uhh… I don't know, but I have thirty new friends on Facebook!"

We are absolutely hammered every day with electronic media. The media tells us what to do, how to be, who to be with, where to spend our money, and that violence is acceptable. And, boy, is it effective. When we watch TV, surf the web, or play on our phones, we go into a mild trance state. When we're in a trance, our subconscious mind is open. This allows images, sounds, and emotions to come right in and attach themselves to the framework of our belief system.

It's extremely difficult not to be subconsciously influenced by media. It is creating and shaping our era. Let's face it: We wouldn't have cell phones without Star Trek. Again, this is neither good nor bad. The question is, "Does this influence produce suffering in our personal world?" This question falls back on the awareness of powerful subconscious emotional charges.

Past Lives

A couple years ago I visited India, a country steeped in the belief of reincarnation. That pervasive belief permeates their entire society, but it's a bit of a two-edged sword. In many instances, a person from India wouldn't bat an eye over things Westerners get completely stressed about. On the other hand, there's not a lot of motivation to get things done. You can always do it in your next life!

I'm not sure about where I stand on reincarnation, but as a healer, it doesn't really matter. Can suffering originate in a past life? Maybe. However, entertaining the notion in subconscious healing provides some enormous benefits. It allows clients to step completely outside of the context and mental framework of their present-day lives and gain an entirely new perspective on their own suffering. It's a way to detach in order to let go.

Let's imagine a young man came to me for a chronic condition that stumped medical doctors. There was no logical explanation for the incredible pain he felt in his body. When he finally arrived at my door, he had been through numerous examinations and a treasure chest full of medications.

Between his work with me and another local alternative health care professional, he was able to heal about 90 percent of the pain over the course of a year. His biggest breakthrough came during a past-life regression session. It was huge. The experience gave him the opportunity to completely reframe the context of his life. It was absolute freedom. Somehow, in the context of that liberty, he was able to free an enormous amount of physical pain. The connection between his stored subconscious emotion and physical pain was released.

Trauma

As previously mentioned, emotional charges can enter the subconscious mind before the critical faculty of the mind is formed in early childhood. Charges can also enter when the critical faculty is let down. This can occur in times of shock and excessive trauma. Examples may include sudden physical injury, surgery, or receiving unexpected shocking news.

When we're stunned, we enter a trance. The critical faculty relaxes, allowing negative emotional information into the subconscious mind. In my experience and for reasons I don't completely understand, traumatic events are not uncommon around the ages of two, fourteen, and twenty.

The Unknown

Sometimes, understanding the cause of an emotional charge can powerfully accelerate subconscious healing. However, every now and then, you can psychologize, analyze, and snoop around for sources of pain without ever finding the source. Luckily, in terms of alleviating subconscious pain, finding the cause isn't absolutely necessary. As we'll learn later on, simply locating the feeling of emotional pain in the body is enough.

Noteworthy Emotional Charge Issues

Before we dive into the process of freeing yourself from subconscious emotional charges and the suffering that accompanies them, let's take a peek at a few charges that many of us experience. The list below includes issues I've had significant firsthand experience in helping people overcome.

Abandonment

Because we're born absolutely dependent, the fear of being left alone is almost instinctual. We're not designed to be isolated. We can't survive by ourselves. In a way, being alone isn't really human. Ironically, we come into this life alone, and we go out alone.

We must also face every life transition by ourselves. Sooner or later, everyone gets kicked out of the nest. To fully mature, we have to stand on our own two feet, and it's not always easy.

"Because we're born absolutely dependent, the fear of being left alone is almost instinctual. "

In regard to abandonment, subconscious emotional charges can be created when there isn't a loving presence to guide us through early phases of childhood. We're too vulnerable, innocent, and ill-equipped to understand or handle initial life transitions on our own. When we're forced to do so, the resulting subconscious fear can cause debilitating circumstances in our adult lives.

Imagine a young man whose father was not around when he was born. In fact, he wasn't around much at all. As a boy, he was frequently told, "Go play outside," and he wasn't allowed to come back inside until suppertime. The time of year or weather conditions didn't matter. He was left unsupervised and alone all day. Fortunately, he lived near some woods, so he wasn't turned directly out onto the streets. Instinctively, he secured hiding places for food and money, in case he wasn't allowed to return back home for good someday, and he grew up tough and afraid.

Two things happened when he matured into an adult. First, he developed incredible independence and resolve. He could take care of himself in any situation. Second, he couldn't bear to be on his own for any extended period of time. The subconscious emotional charges he carried from his youth produced absolute terror when he was left alone.

He eventually became very successful in a business that required extensive travel. However, he had multiple lovers stashed in different cities, just like the food and money he used to hide in the woods. He was never alone.

Unworthiness, Rejection, and Insecurity—Lack of Love

This one is a biggie because most of us get at least one helping of it. None of us get enough love. It's just not possible. We enter this world completely dependent on someone else, physically and emotionally, twenty-four hours a day, seven days a week. There hasn't been one set of parents in all of human history who has had the ability to satisfy the insatiable needs of an infant. It truly does take a village to raise a child. Add another kid or two into the mix and forget about it.

We grow up competing for love and attention. We'll do anything to get it. We'll be perfect, troublesome, depressed, sick, hurt, or even a superstar. Throwing some parental neglect and abuse on this ravenous need can produce some deeply seated emotional charges.

Unworthiness can be expressed in a variety of ways but tends to gravitate into two extremes: under- and overachieving. This is not only limited to jobs and careers; it can also spill into relationships. Examples may include the workaholic executive who is always vying for supervisor approval or someone who continues to sabotage seemingly healthy intimate relationships. The underlying subconscious message is the same: I'm not worthy.

The Broken Heart — Grieving

Life is a never-ending parade of hanging on and letting go, letting go and hanging on. In fact, to be happy, healthy, and well adjusted, we need to learn the art of letting go. It's imperative to learn how to grieve.

At some point in life, we discover that being human means being brokenhearted. We love and become emotionally attached. When we lose someone or something we love, it can be very painful, resulting in deep grief. It can be over anything: a person, an object, our innocence, a pet, a part of ourselves, circumstances, etc. It arrives in waves over a period of time. Just when you think you're finished, along comes another wave. It can take weeks, months, or even years. The timeframe is different for everyone.

“To be happy, healthy, and well adjusted, we need to learn the art of letting go.”

A variety of emotions are experienced during the grieving process. They may include shock, denial, sadness, loneliness, depression, anger, and at some point hopefully, acceptance,a decision to move on, forgiveness, and peace. Problems arise when we don’t allow ourselves to grieve. This can be due to lack of awareness or flat-out refusal to experience difficult emotions. In either case, the suppression of emotion will create an emotional charge in the subconscious mind. The charge will then continue to repeat itself over time, until it’s experienced.

There are times when life can be difficult. Grieving is one of them. In the framework of spiritual growth, grieving is not a time to shy away from the work that needs to done. It’s a time to allow yourself, as gently as possible, to feel the feelings of loss and let them go. It’s an act of honoring yourself and what you have lost.

Physical Illness — Cancer

Subconscious emotional charges can be felt in the body.

Perhaps you feel stress in your stomach, rage in your head, or sadness in your heart. The charges need an outlet, a way to be experienced and released. When suppressed, they can find that outlet through the body in the form of illness, instead of just through life circumstance.

I’ve worked extensively with cancer patients over the past decade, and both of my parents died from the disease. When someone is diagnosed with cancer or any serious physical illness, everything changes… immediately. Decisions on treatment must be made, family members need to be notified, finances become a concern, and questions of death come to the forefront. Priorities come sharply into focus. Emotions surface, sometimes powerfully.

My job has been to assist patients through the subconscious emotional mine fields that come up during cancer treatment. Some want healing, both emotionally and physically. Some want to connect to their God. There are those that I must prepare for death. At times, I’ve seen nothing short of miracles; at other times, I’ve seen people go very quickly.

Working with cancer patients has shown me the resilience and frailty of life. I'm constantly reminded who is in charge: the Divine. Cancer patients are some of the most remarkable people I've ever met. Survivors of a life-threatening illness will be forever changed. Whether they seek therapy or not, those who recover go through an enormous internal transformation. With a massive amount of courage, they face their darkest demons.

"Many of us have emotional charges concerning unworthiness because there wasn't enough love to go around in childhood, even in the midst of a loving family."

I've noticed one common subconscious theme among several cancer patients. It concerns issues around unworthiness and lack of love, but at a whole new level. As mentioned above, many of us have subconscious emotional charges concerning unworthiness because there wasn't enough love to go around in childhood, even in the midst of a loving family. Some cancer patients have experienced a lack of love not only through circumstances, but also through actuality.

They simply were not loved as children; they weren't wanted. The realization of that truth can be a very bitter pill to swallow.

I've also noted parallels of those who survive cancer. Many have a very positive attitude and a willingness to be involved in their recovery. Some see cancer as the best thing that ever happened to them because it brings wonderful new people into their life. Others view it as a wake-up call, teaching them what is really important. Many learn about healing, and some connect to the Divine.

I've seen too much success to doubt the powerful connection between the subconscious mind and illness in the physical body, especially when it comes to cancer. When forming a comprehensive plan for the treatment of a life-threatening disease, it's an extremely good idea to incorporate the healing of a potential contributing factor: the subconscious mind.

Exercises for Healing Emotional Blockages and Harnessing the Power of the Subconscious Mind

Preparation for Clearing Emotional Charges From the Mind

Take a few moments to sit and contemplate. You can journal, if that's your style. Think of the most significant difficulties you've had in your life or troubles you may be experiencing now. If you hit a snag, you needn't look far. Most of us don't need to search any farther than our own immediate families. Can you see any repeating patterns? Are you aware of any emotion that has been repressed?

Healing the Mind Exercise 1:
The Curious Mind—Feel the Charge until It's Gone

Note: Before diving into specific subconscious healing techniques, it's important to develop the proper state of mind in regard to emotional charges. It's referred to as the Curious Mind. The Curious Mind naturally cultivates a certain level of necessary detachment between you and the emotional charge. All of a sudden, it's not your anger, hurt, depression, etc.; it's only an emotion you're experiencing. In other words, your emotional charges are not you, but are only something that's happening to you. As you will experience, this creates a beneficial space between you and the emotion—a very helpful gap that enhances the release process.

Try This: Start with a clear-cut intention to heal a specific subconscious emotional charge completely. Sit or lie in a comfortable position and enter a deeply relaxed state. Next, feel the emotional charge in your body, as related to the area of life you're trying to heal. It may be rage in your head, heaviness in your heart, tension in your stomach, etc. It should come readily. If it doesn't, just pretend you can feel it (remember, imagination is real, so pretending works). If you get bombarded with mind chatter or distracting thoughts, just let them come and go, like watching clouds drifting by in the sky.

Next, simply pay attention to the charge as you experience it. Be curious. Say to yourself something like, "Oh, isn't that interesting. I feel rage in my head. How curious." As you focus on that experience, you'll begin to feel a weakening, softening, or letting go of the emotion.

Now, here's where everyone begins to differ, so you'll need to develop your own method of experiencing and releasing charges. You may feel the emotion as it begins to spin or swirl and move out of your body. It may start to dissolve or let go. It may create interesting sensations in your body like tingling, warmth, or deep peace. Whatever happens, keep your focus on what's taking place. Don't wallow in the emotion, but stay with it until it's gone.

That's it! Subconscious emotional charge gone. If it returns, it will come back with less intensity and less frequency. If that happens, just go through the process again. It's like lighting a log on a fire: Keep burning it until it's nothing but ash.

Healing the Mind Exercise 2:
Flower or Sacred Object—General Emotional Release

Sometimes when using the Curious Mind approach, charges get stuck. They just won't go away. You end up flailing and wallowing in unwanted emotion. It can be a very frustrating experience. However, we have help. After all, it is the subconscious mind, and we can imagine whatever we want to come to our aid. The Flower or Sacred Object is a technique I've used countless times on clients, with immediate, successful results. Here's how it works:

Try This: While relaxed and feeling your subconscious emotional charge, imagine a beautiful flower or some sacred object hovering above you. Let it be big and powerful. Feel its presence. Recognize the fact that it has enormous healing powers and actually pulls emotional charges out of your body. At this point, you don't have to do anything! Let the flower or sacred object do all the work by pulling the charge out of your body. Keep your focus on the process until it's finished.

Healing the Mind Exercise 3:
Self-Forgiveness—Unloved and Unworthy (It's Not Your Fault)

This is probably the most powerful technique I've come across. It can mend and fill an eternal abyss of feeling unloved. Reaching very deep, it can transform an entire life of suffering.

Try This: Settle into a meditative state. Imagine yourself going back into time, just before your conception. You have no body, and you are nothing but energy. You're about to be conceived in an act of absolute unconditional love between two people who are the embodiment of that love. Sink deeply into the experience that as you grow in the womb, your arrival is eagerly anticipated. People are excited! You are wanted!

Imagine you are born into the arms of a perfect parent. It can be someone you make up or an actual person you know. Feel unconditional love exuding from them like sunshine.

Have them look deeply into your eyes, connecting heart to heart, and hear them say, “It’s not your fault. It has never been your fault. It will never be your fault.” Sink into the experience fully. Don’t just hear the words, but receive the impact of that unconditional love into your psyche. Take it in deep because this sensation can lead to a very positive, life-changing transformation.

Healing the Mind Exercise 4:
Forgiveness—It’s Not Their Fault

Forgiveness, in this context of subconscious healing, is a little different than the human virtue. I discuss the importance of that in my book *The Subconscious, The Divine and Me*. Here, let’s focus on a technique to release a subconscious emotional charge.

Try This: Get relaxed and think of someone whom you need to forgive. Feel the emotional charge when you think of that person, and take note of where you feel it in your body.

Now, imagine yourself standing somewhere: the woods, a beach, at home, or anyplace that comes to mind. Next, imagine a small child approaching you, maybe three or four years old. As the child nears, recognize that s/he is the person you need to forgive. Observe the child as vulnerable, innocent, and at the mercy of life, the same way you were when you were that age. Be aware that the child has been a victim of his/her parents’ emotional charges and understand that it’s not the child’s fault. Sink into that experience and feel it and then let your own pain go. Use an above mentioned technique to release your charge or just send it back to the child. Then let the child go to dissolve into the cosmos. Or, you can pick the child up and say, “It’s not your fault.” Feel it deeply and remain focused on the process until it has run its course.

The piece on the subconscious mind was excerpted with permission from The Subconscious, the Divine and Me: A Spiritual Guide for the Day-to-Day Pilgrim by Joseph Drumheller.

Scientist says human consciousness comes from another dimension Jun 29th 2025

By ELLYN LAPOINTE FOR DAILYMAIL.COM

A baffling new **theory to explain human consciousness** has suggested it comes from hidden dimensions and is not just brain activity. A physicist claimed that we plug in to these invisible planes of the universe when making art, practicing science, pondering philosophy or dreaming, and this could explain the phenomenon that has evaded scientific understanding for centuries.

Michael Pravica, a professor of physics at the University of **Nevada**, **Las Vegas**, has based the wild idea on hyperdimensionality, the idea that the universe is made up of more dimensions than just the four we perceive: height, length width and time. But his theory is highly controversial, with one scientist saying that the cornerstone of Pravica's theory 'borders on science fiction.'

© Getty Images/Science Photo Libra

Physicist Michael Pravica believes that human consciousness transcend the physical world and move between hidden dimensions

'The sheer fact that we can conceive of higher dimensions than four within our mind, within our mathematics, is a gift... it's something that transcends biology.'
Scientists have been attempting to explain human consciousness and its origins for hundreds of years - and the theories run the gamut.

One leading theory suggests that consciousness is related to how much information is integrated between the different parts of the brain. The more information is connected and integrated, the more conscious a being is thought to be.

Another posits that conscious mental states are driven by top-down signaling in the brain. Top-down signaling refers to the process by which higher-level brain regions send information, expectations or context to lower-level brain regions.

But Pravica's theory ventures outside the realm of neuroscience and into theoretical physics.

He suggested that in moments of heightened awareness, like when we enter a dream state or use our brains for deeply creative or intellectual tasks, our consciousness could transcend our physical dimension and enter a higher plane.

In these moments, our consciousness syncs with hidden dimensions and receives a flood of inspiration, Pravica said.

To better understand the controversial theory, consider the following scenario.

Imagine you're a two-dimensional being living in a two-dimensional world, like a character in a comic book. Now, imagine that a sphere passes through your plane of view.

The sphere would look like a dot that grows into a larger and larger circle as it comes closer, then gradually shrinks until it's out of view. You would have no way of knowing that it's actually a three-dimensional shape.

Pravica sees us as a version of these 2D characters. Although we exist in a four-dimensional world, we can only perceive matter and energy that is of those four dimensions, just like how beings in a 2D world cannot perceive a 3D object.

Thus, the limitations of our world prevent us from detecting higher dimensions that could, in theory, exist all around us.

This is the foundation of hyperdimensionality - the idea that the universe is made up of many dimensions, some of which are hidden because they are beyond the reach of our physical realm.

Hyperdimensionality ties into string theory, which states that reality is made up of infinitely small vibrating strings that are smaller than atoms, electrons or quarks.

As the strings vibrate, twist and fold, they produce effects in multiple unseen dimensions that give rise to all the particles and forces that we can observe, from particle physics to gravity.

'String theory is essentially a theory of hyperdimensionality,' Pravica said. 'It's looking at how the universe is put together on a sub-quantum scale.'


© Getty Images

Pravica believes that our brains can tap into higher dimensions when in a dream state or performing deeply creating or intellectual tasks

Although we can observe the effects that these vibrating strings have on the physics of our dimension, we can't observe the hidden dimensions that they're vibrating in.

That is - we can't *physically* observe them.

But our consciousness may be able to tap into them, Pravica says.

Hyperdimensionality and string theory are widely accepted by physicists, but Pravica's idea of their relationship with consciousness is more controversial - especially because it blurs the lines between science and spirituality.

As an Orthodox Christian with a Ph.D. from Harvard, Pravica has found hyperdimensionality to be a way to bridge his scientific background with his religious beliefs.

For example, he believes Jesus may be a hyperdimensional being.

'According to the Bible, Jesus ascended into heaven 40 days after being on Earth. How do you ascend into heaven if you're a four-dimensional creature?' Pravica asked.

But being hyperdimensional could, theoretically, have allowed Jesus to move between our world and heaven - which may be a world of higher or infinite dimensions, he said.

Pravica's theory is based on a 'God of the gaps' perspective, where gaps in scientific knowledge are explained by divine intervention, said Stephen Holler, associate professor of physics at Fordham University.

He believes that this type of thinking is insufficient, and hampers the scientific inquiry needed to truly understand and explain ineffable phenomena like human consciousness.

'It's a poor explanation mechanism that arguably stifles the inquisitive nature required for good science and teaches that it's not okay to say, 'I don't know,'' Holler told Popular Mechanics.

He points out that our ability to mathematically manipulate higher dimensions is not proof that they actually exist, or that our consciousness can interact with them.

What's more, exploring these higher dimensions is impossible due to the limitations of our current technological capabilities.

Not even the most powerful particle accelerator in the world - the Large Hadron Collider (LHC) at CERN - can provide real proof that these dimensions exist.

The LHC smashes particles together at incredibly high speeds - up to the speed of light.

This allows physicists to study the fundamental building blocks of matter and energy and access infinitesimally small dimensions - even smaller than a single proton.

But even the LHC isn't able to reveal the high-dimensional strings that quantum physics predicts. To get that granular, physicists would need a much more powerful collider.

Without that concrete evidence, Holler says that hyperdimensionality 'borders on science fiction.'

But Pravica is optimistic that such technology could exist within his children's lifetime.

Until then, he will continue to support hyperdimensionality and his theory of how it relates to our consciousness.

'I see no point otherwise,' he said. 'Why study? Why live?'

A baffling new theory to explain human consciousness has suggested it comes from hidden dimensions and is not just brain activity.

A physicist claimed that we plug in to these invisible planes of the universe when making art, practicing science, pondering philosophy or dreaming, and this could explain the phenomenon that has evaded scientific understanding for centuries.

To better understand the controversial theory, consider the following scenario.

Imagine you're a two-dimensional being living in a two-dimensional world, like a character in a comic book. Now, imagine that a sphere passes through your plane of view. The sphere would look like a dot that grows into a larger and larger circle as it comes closer, then gradually shrinks until it's out of view. You would have no way of knowing that it's actually a three-dimensional shape.

Pravica sees us as a version of these 2D characters. Although we exist in a four-dimensional world, we can only perceive matter and energy that is of those four dimensions, just like how beings in a 2D world cannot perceive a 3D object.

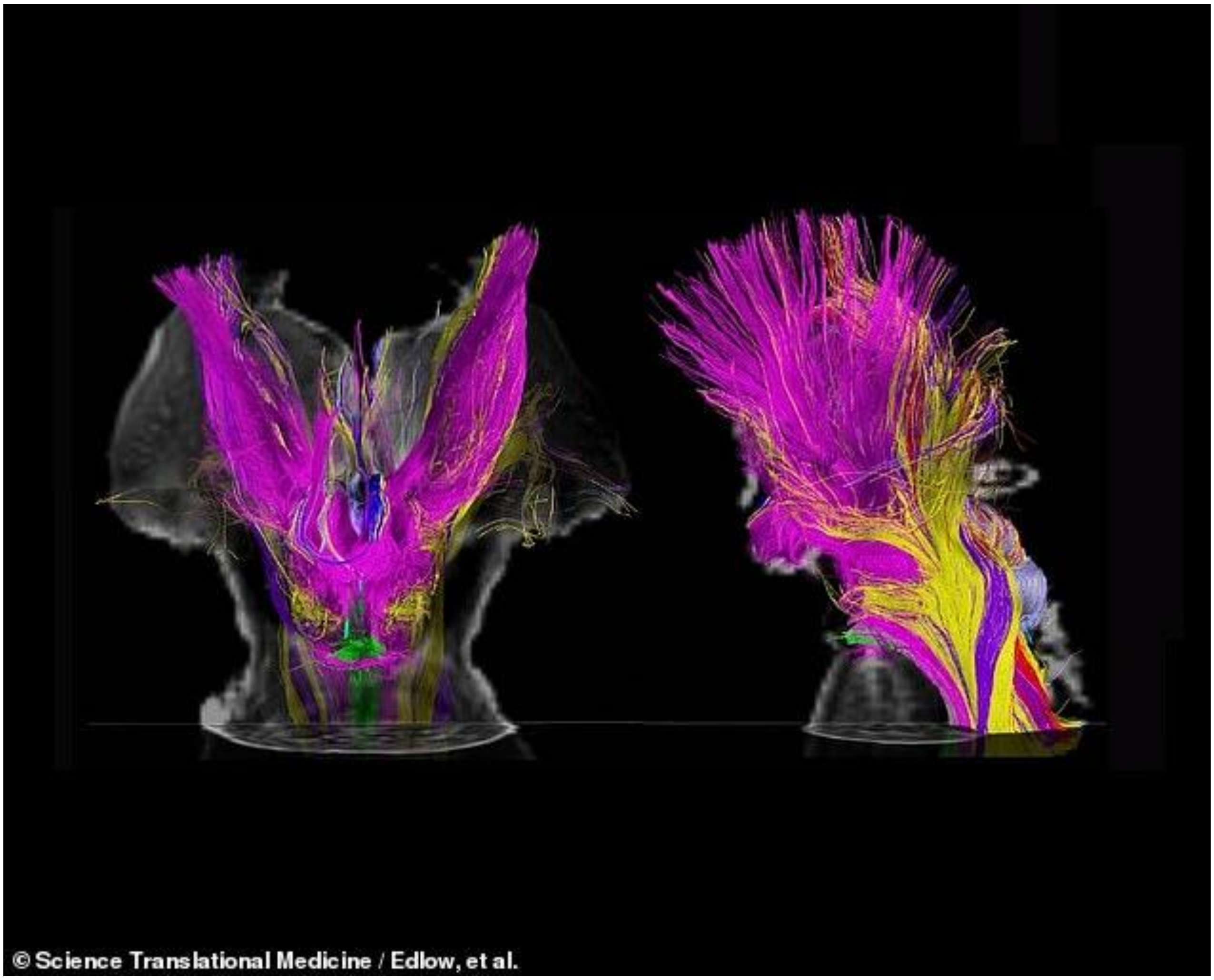
Thus, the limitations of our world prevent us from detecting higher dimensions that could, in theory, exist all around us. This is the foundation of hyperdimensionality - the idea that the universe is made up of many dimensions, some of which are hidden because they are beyond the reach of our physical realm.

MRI scans reveal the STUNNING stages of consciousness in the brain, Jun 30th 2025

Enduring questions over which part of the brain helps produce that feeling of being 'awake' have been answered, thanks to stunningly detailed new brain imagery.

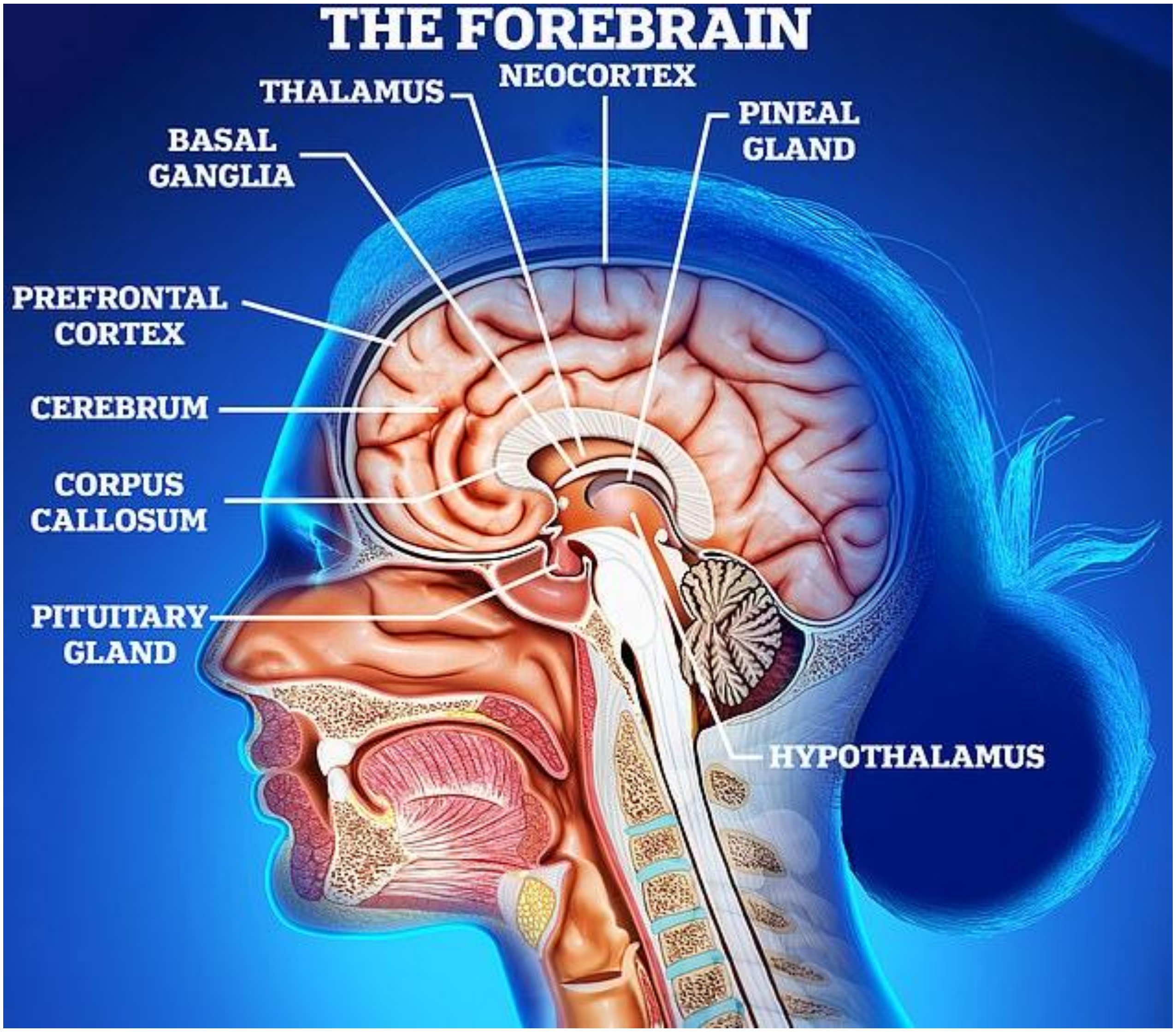
Researchers' new high-resolution brain scans allowed them to see brain connections at a granular 'submillimeter' level — meaning down to a tiny 3/100ths of an inch. The images were then used to map a neural network of previously unseen pathways in the brain, called the 'default ascending arousal network' or dAAN, which they now theorize is the core region that helps humans sustain wakeful consciousness.

In recent years, the neuroscientists studying consciousness have divided this curious mystery of how the human brain is self-aware into two sub-categories: 'arousal' (wakefulness) and 'awareness' (the subjective experience of being alive). The researchers hope their work exploring this dAAN pathway will help develop new treatments for patients with comas, or other conditions that hinge on wakefulness.



© Science Translational Medicine / Edlow, et al.

Researchers' new high-resolution brain scans (above) allowed them to see brain connections at a granular 'submillimeter' level - meaning down to a tiny 3/100ths of an inch



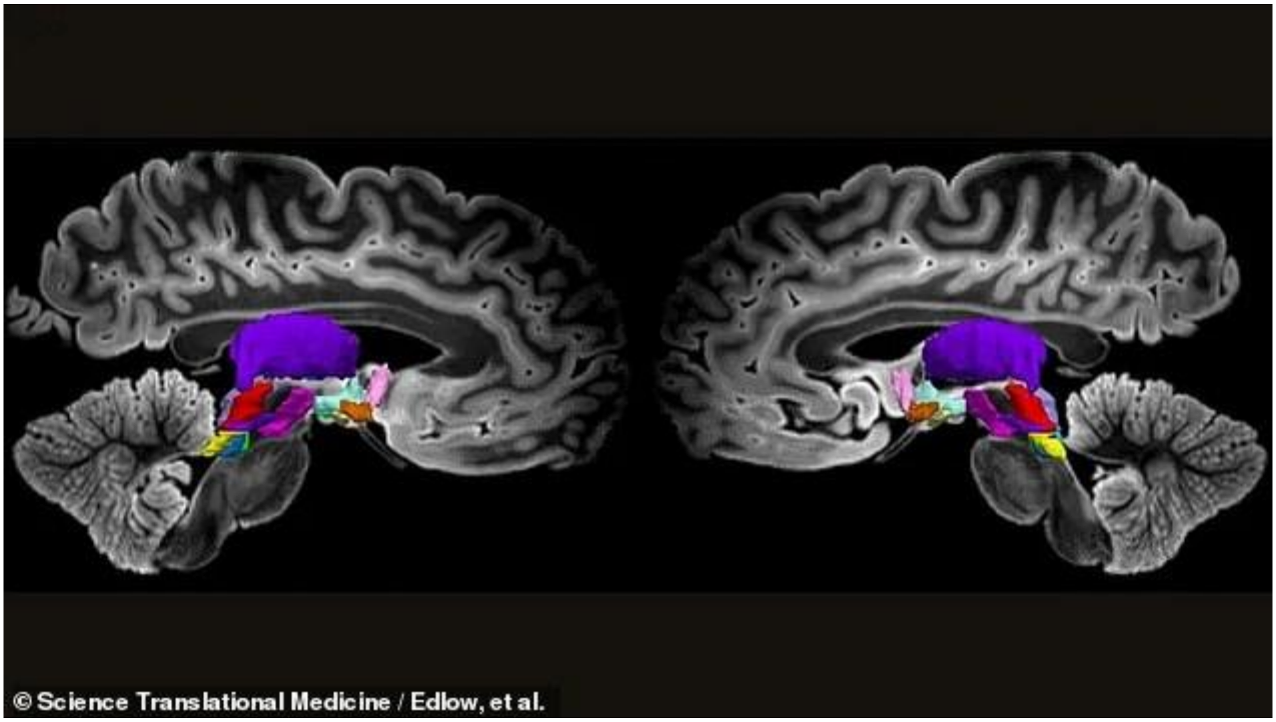
The regions they mapped were centered around the base of the brain where it connects to the spine: including the brainstem, thalamus, hypothalamus, basal forebrain, and cerebral cortex

'Our goal was to map a human brain network that is critical to consciousness,' lead author Dr. Brian Edlow said, 'to provide clinicians with better tools to detect, predict, and promote recovery of consciousness in patients with severe brain injuries.'

Dr. Edlow's colleague and the new study's senior author Dr. Hannah Kinney, specified that 'a broad range of neurological disorders' associated with altered states of consciousness could also benefit from the new research.

Dr. Kinney, a professor emerita at Harvard Medical School and Boston Children's Hospital, noted that the new neural maps may also one day help patients coping with seizures and sudden infant death syndrome, as well as those recovering from comas.

Dr. Nicholas Schiff, a professor of neurology and neuroscience at Weill Cornell Medicine — who was not involved in the new research — told **Live Science** that the team's neural mapping was an impressive new achievement.
'It's a beautiful study,' Dr. Schiff said. 'It's a map of everything.'


© Science Translational Medicine / Edlow, et al.

Above, an MRI scan of the major nodes of the human brain that the new study's authors argue are key to keeping humans awake. The nodes have been color-coded for their neural mapping

The study's authors hope their work exploring the conscious brain's 'wakefulness' pathways will help develop new treatments for coma patients, as well as those coping with seizures

Using a technique called 'ex vivo magnetic resonance imaging' (MRI), a slower form of MRI scanning conducted on dead or surgically removed tissue, the researchers were able to grab detailed images of the dAAN pathways linked to consciousness.

The regions they mapped were all tightly centered around the center or base of the brain, close to where it connects to the spine: including the brainstem, thalamus, hypothalamus, basal forebrain, and cerebral cortex.

Delving inside the mind: Incredible graphics reveal what each section of your BRAIN does - with more than 70,000 thoughts processed every single day

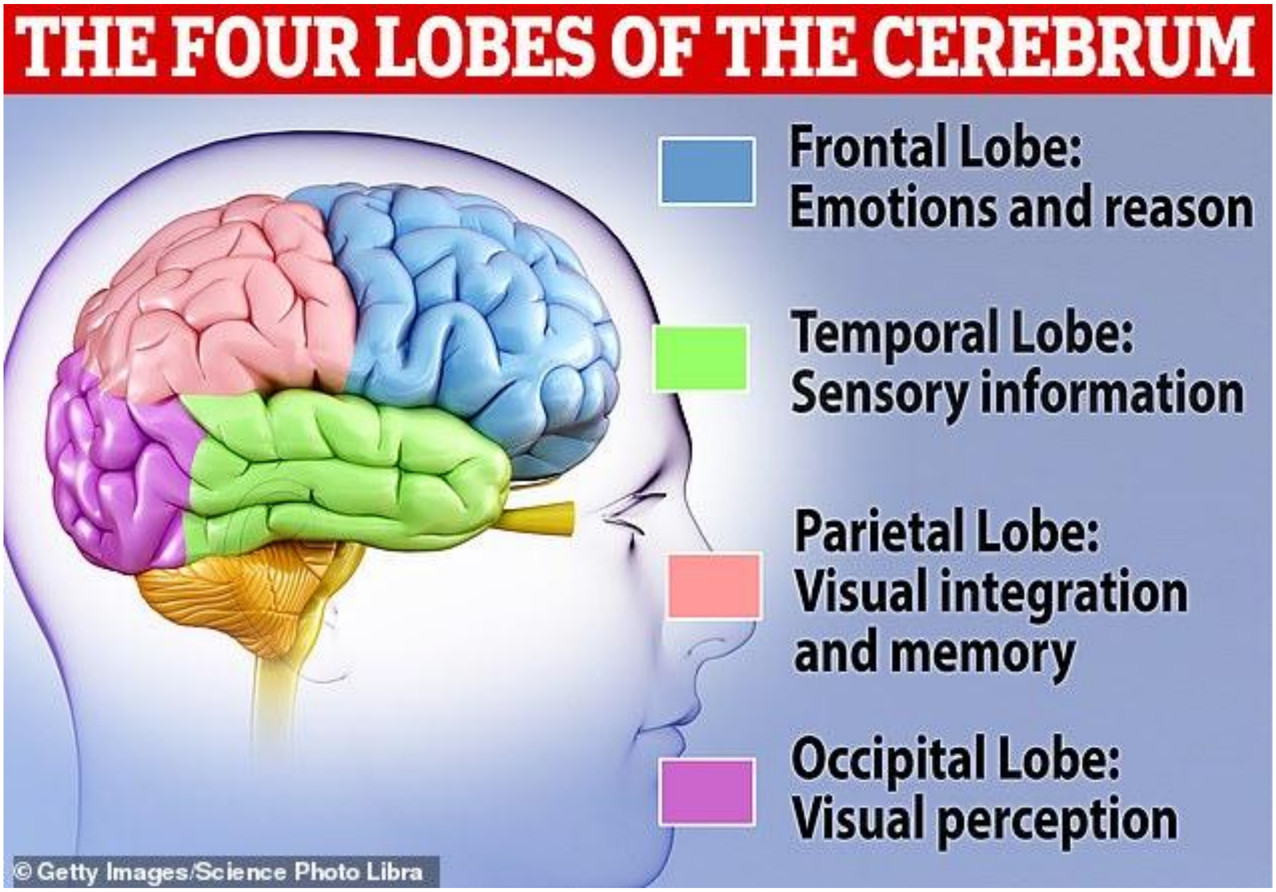
With an estimated 86 billion neurons, the 3lb organ between your ears may be one of the most complex objects in the universe.

All that complexity allows the brain to process more than 70,000 thoughts every day - but creates quite a headache for neuroscientists.

Research has identified hundreds of different brain areas and, the closer scientists look, the more structures seem to appear.

Incredible graphics now reveal the weird and wonderful functions of these different areas - from the lobes that create your personality to the mysterious origins of language.

So, if you've ever wondered just what is going on between your ears, read on to find out how the shape of the brain shapes your mind - and what happens when it goes wrong.



At the simplest level we can divided the brain into lobes which are responsible for emotions and reason, sensory information, memory, and visual perception

At a very simple level, the brain has four main areas, each with its own component parts.

Around the outside of the brain there is the Cerebrum, one of the biggest structures which includes the wrinkly Cerebral Cortex.

Peeling this outer layer back, we reach the Forebrain, a complex set of different glands and brain regions where we find the key control centres of the body.

Go another layer deeper and, nestled at the very core of the brain is the Midbrain.

Finally, connecting the rest of the parts to the nervous system, there is the Hindbrain, believed to be the most evolutionarily ancient part of the human brain.

The Cerebrum

If you imagine a brain, what you are envisioning is most likely the cerebrum.

The pinkish wrinkly mass of folded grey matter makes up around 85 per cent of the entire human brain and is where most of our higher mental functions occur.

Planning, language, and thought - all the things that make us distinctly human - have their origins somewhere in this region of the brain.

The very outer layer of the Cerebrum, that pink layer normally seen in pictures, is called the Cerebral Cortex and is typically divided into four lobes: the Frontal, Parietal, Temporal and Occipital.

The Frontal Lobe is believed to be where our higher executive functions take place such as planning, reasoning, and controlling our emotions.

In cases where the Frontal Lobe is damaged, we often observe sudden and dramatic changes to someone's personality.

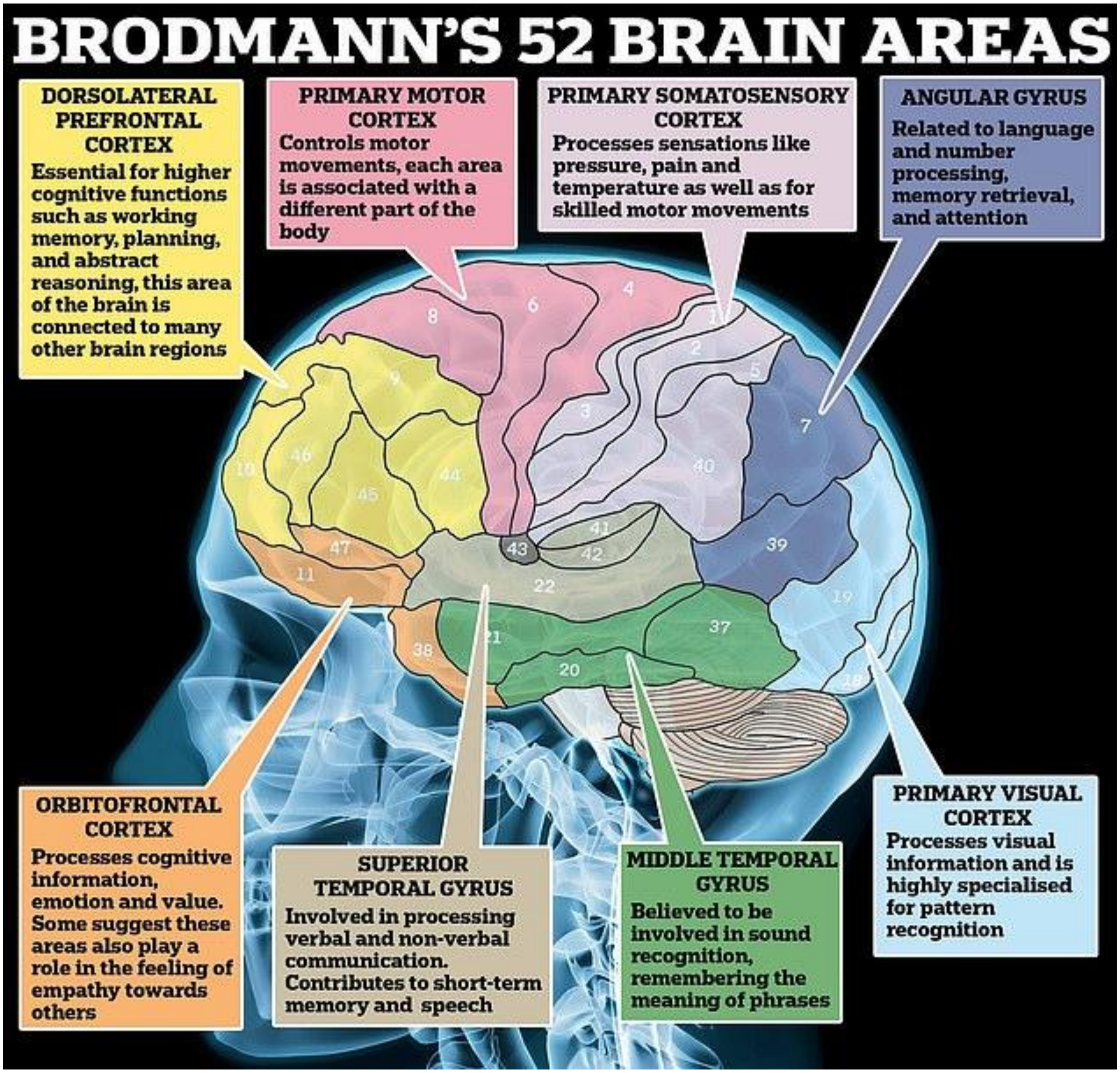
The Parietal Lobe is believed to be responsible for integrating sensory information like touch, temperature, and pain.

The Temporal Lobe is dedicated to processing sensory information, particularly hearing and recognizing language.

Finally, the Occipital Lobe is the brain's visual processing centre.

This region receives signals from the eyes which it sends to several different areas to process information such as depth, location, and the identity of objects.

However, scientists haven't stopped with just the four big lobes, as the cerebral cortex can be further divided into 52 regions known as Brodmann's areas.



Brodmann assigned each of the brain regions a number, some of which have been associated with functions such as memory recall, emotional control, and abstract reasoning

Published in 1909, Korbinian Brodmann's groundbreaking analysis of the brain can still be found in neurology textbooks and on classroom posters to this day.

Using a specialized microscope, Brodmann painstakingly analyzed the entire surface of the Cerebral Cortex on cellular structure alone.

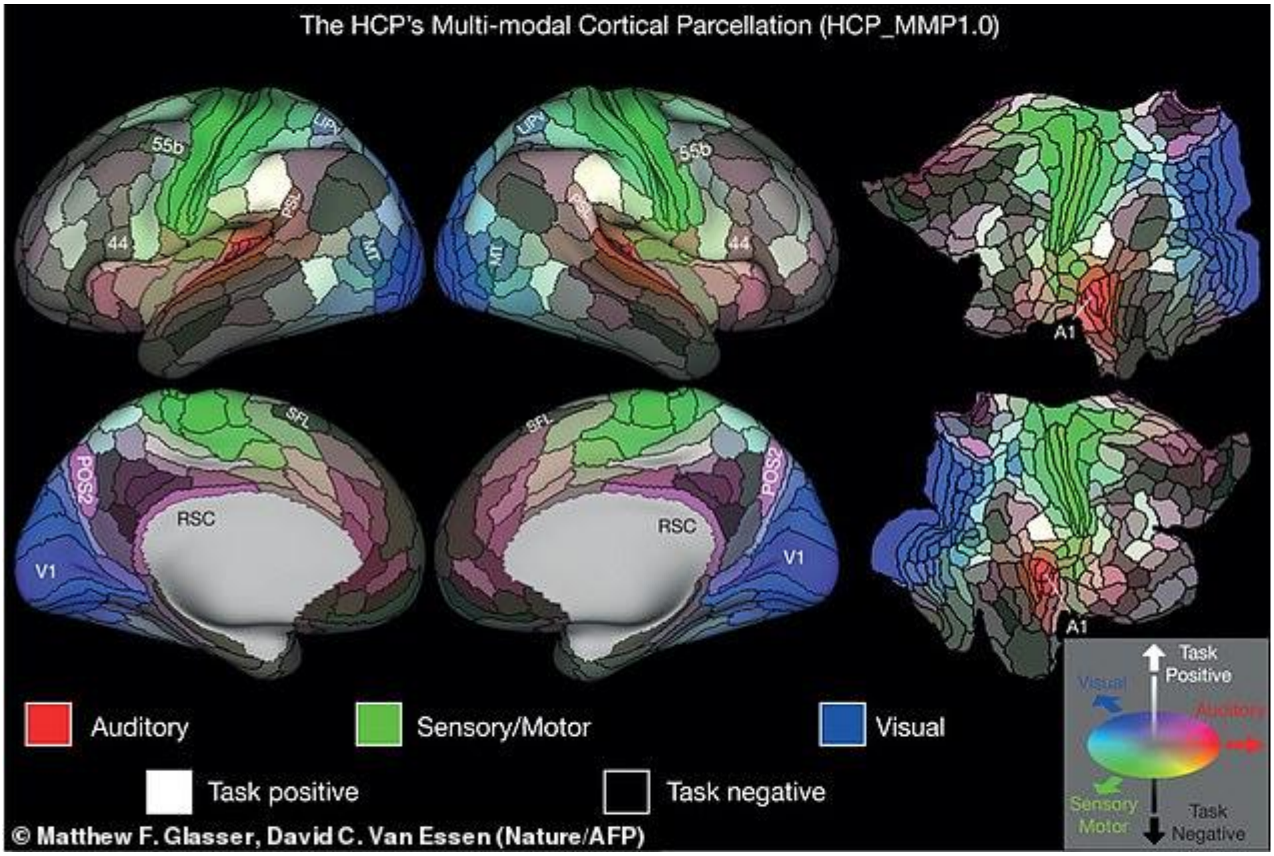
After a decade of effort, Brodmann produced the most detailed map of the Cerebral Cortex yet produced, assigning each region a different number.

Over time these areas have been widely used to link brain regions with specific functions, such as area four: the primary motor cortex.

This region of the Cerebral Cortex is believed to control motor movements such as moving the hands and face as well as breathing and voluntary blinking.

Brodmann's areas have also been mapped to functions such as processing numbers, planning, and processing emotions.

Of course, the complexity doesn't stop there as scientists now believe the **Cortex has at least 180 distinct regions** important for language, perception, consciousness, and attention.



The HCP's Multi-modal Cortical Parcellation (HCP_MMP1.0)

© Matthew F. Glasser, David C. Van Essen (Nature/AFP)

Scientists have now revealed that there are more than 180 distinct regions in the Cerebral Cortex, more than tripling the number of areas found by Brodmann

The Forebrain

The Cerebrum is part of a structure called the Forebrain, which makes up the majority of the brain and controls most of our important bodily functions.

In addition to the Cerebrum and Cerebral Cortex, the Forebrain also contains structures such as the Thalamus, Hypothalamus, Corpus Callosum, and Pituitary gland.

The Thalamus functions like a 'relay station' for the brain's sensory information processing, linking up the spinal cord to other parts of the brain.

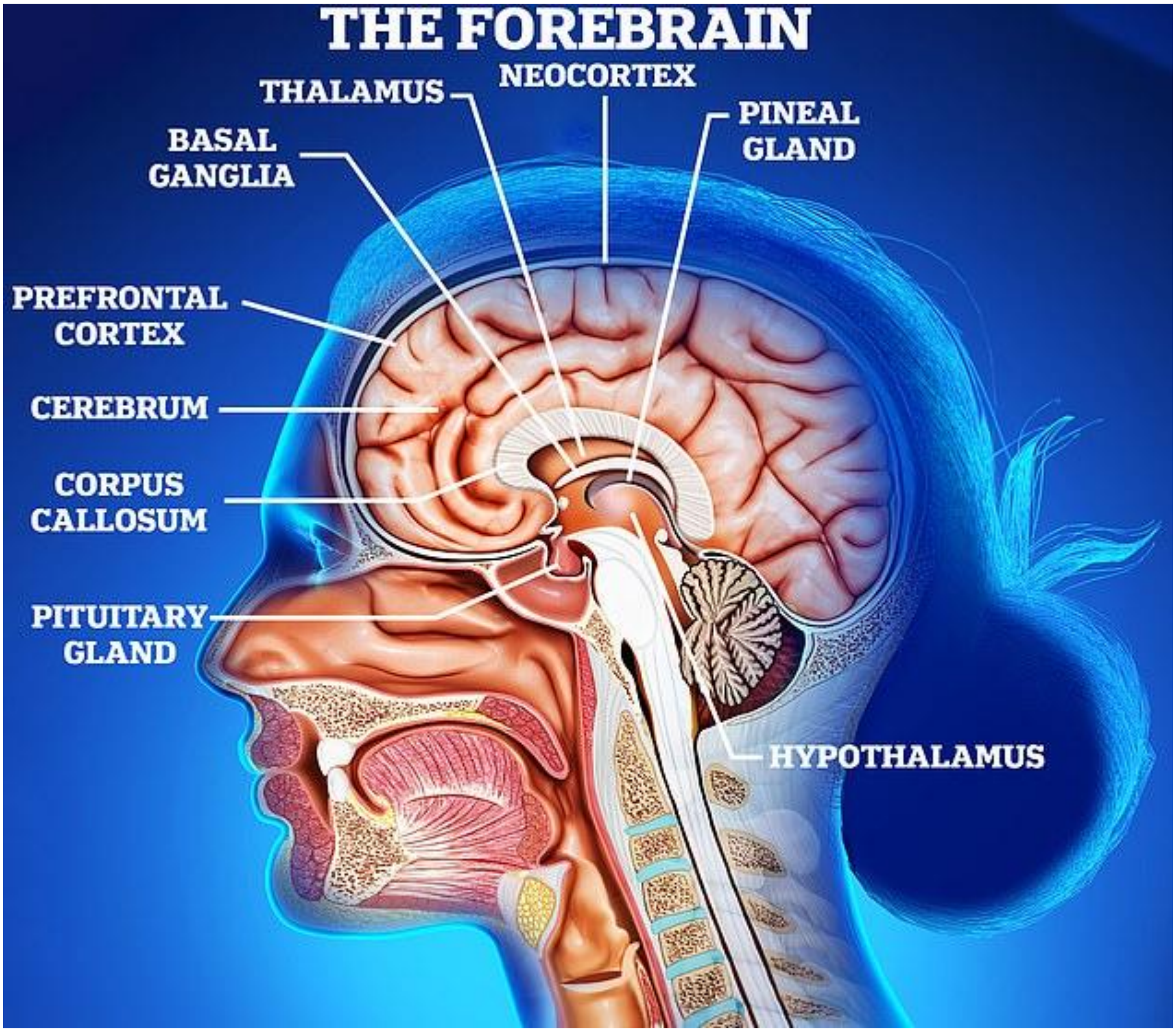
The Hypothalamus, however, which sits just beneath the Thalamus, controls our 'autonomic motor system'.

This means it is responsible for regulating everything from heart rate and breathing to digestion and sexual arousal.

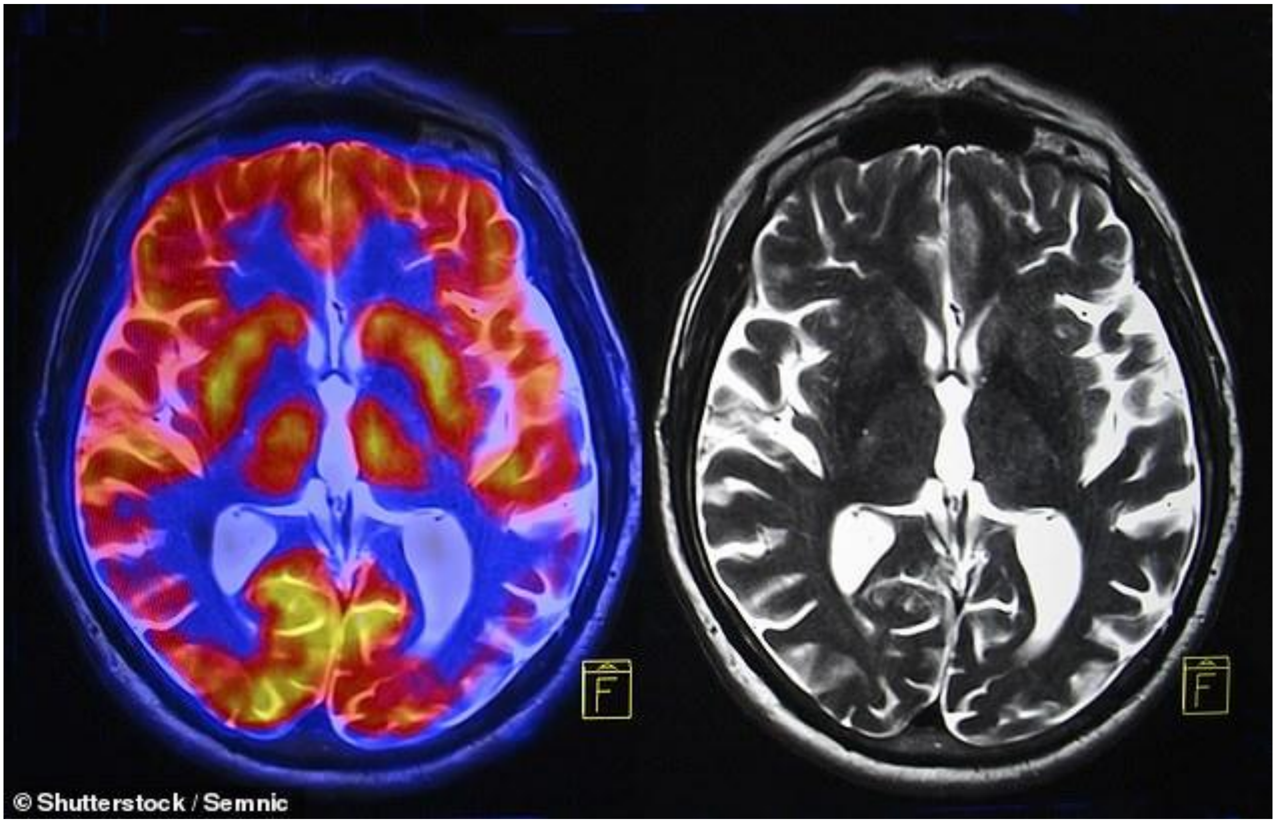
Tucked in between the two lobes of the Thalamus is an even smaller but vital region called the pineal gland.

Called the 'seat of the soul' by philosopher Renee Descartes, this small gland was once believed to be the point at which the immaterial soul interfaced with the body.

Although its function is more mundane it is no less vital as the Pineal Gland controls our sleep patterns by secreting the hormone melatonin.



The Forebrain is the largest section of the human brain and is responsible for most of our higher thought, injuries to areas like the Corpus Callosum can have serious consequences



© Shutterstock / Semnic

This scan of the brain shows the inner areas of the Midbrain and Hindbrain, around the outside you can see the wrinkly layer of the Cerebral Cortex where higher thought processes take place

Another important section in the Forebrain is a region called the Corpus Collosum which is the band of white matter which connects the left and right hemispheres.

Made up of more than 200 million nerve fibers the Corpus Collosum is the largest connective pathway in the brain and keeps growing until we reach our 30s.

In some cases epilepsy patients will have this region severed as a last resort to stop seizures spreading throughout the brain.

But these operations, called Corpus Callosotomy, sometimes create so called 'split brain cases'.

Damage to the Corpus Callosum is associated with 'alien hand syndrome', or 'Dr Strangelove Syndrome' as it is often called in reference to the 1964 film Dr. Strangelove.

This can lead patients to lose control of their own hands, with their limbs acting as if controlled by another consciousness.

Psychologists Michael Gazzinga and Roger Sperry described the case of a patient whose left hand tried to strike his wife while the right fought to stop it.

In 2011 a woman named Karen Byrne reported that she was frequently attacked by her own hand after undergoing a Corpus Callosotomy.

The Midbrain

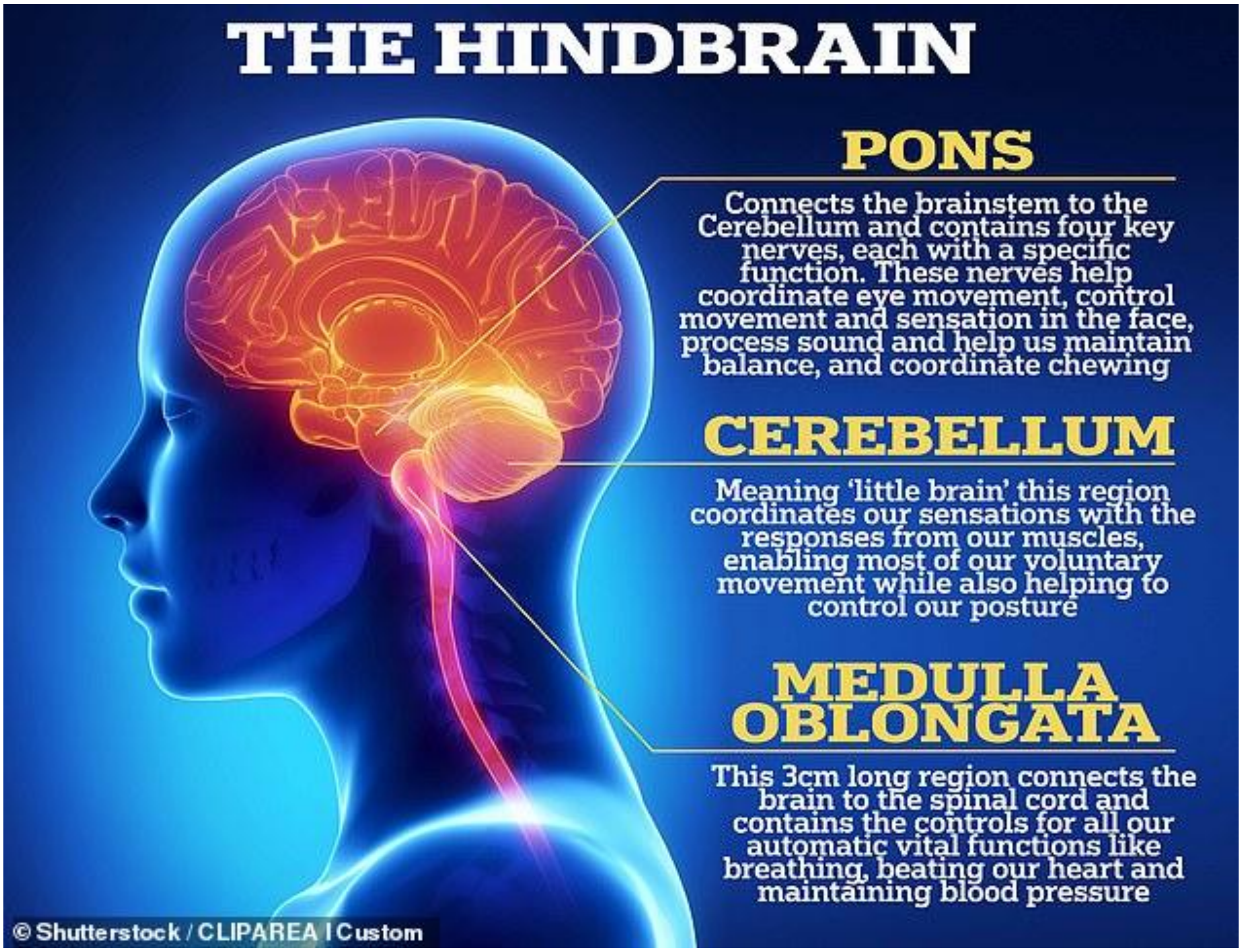
Move another layer deeper and you find the Midbrain, the smallest structure of the brainstem at just over half an inch (1.5cm) thick.

However, it is indispensable for the movement of signals between the nervous system in the body and the processing centres of the brain.

The Midbrain plays an important role in our ability to understand where we are in space as well as controlling our movement.

Scientists also believe that parts of the Midbrain might play an important role in both consciousness and sleep.

The Hindbrain



The Hindbrain is the most ancient structure that makes up the human brain, it is responsible for overseeing our basic survival functions but regions like the Cerebellum have also been linked to functions like abstract thought and creativity

Finally, we arrive at the deepest and oldest structure of the human brain: the Hindbrain.

This region, which makes up most of the brainstem, connects the spine to the brain and is where we find our most basic functions that are essential for life.

At the very top of the Hindbrain is the Pons, which takes its name from the Latin word for bridge, and acts as the interface between the rest of the Hindbrain and the Cerebral Cortex.

The Pons contains four nerves which each have a unique function.

These nerves contribute to eye movement, sensations in the face, balance and hearing, and chewing.

The lowest part of the Hindbrain is the Medulla Oblongata, a one-inch-long lump of white and grey matter which plays a vital role in the body's automatic functions.

Without the Medulla Oblongata, you would be unable to control your heart rate, blood pressure, breathing, or even sneezing.

At last, we arrive at the Cerebellum, meaning 'little brain', this is the largest part of the Hindbrain and is part of the evolutionary ancient structure sometimes called the 'lizard brain'.

Just like the Cerebrum, the Cerebellum has a wrinkly outer layer of grey matter and is divided into two hemispheres.

The Cerebellum is so densely packed that it contains 80 per cent of all the neurons in the brain.

Scientists believed it originally evolved to facilitate the smooth execution of motor functions.

However, it is also believed to play a big role in abstract thinking and several studies link insight and creativity to the Cerebellum.

Neuroimaging has allowed us to peer deeper into the human brain than ever before, but the question of how to explain the brain's functions is still one scientists are grappling with

How do we know what the brain does?

Since the Ancient Greeks, thinkers and philosophers alike have been puzzling over the function of the brain.

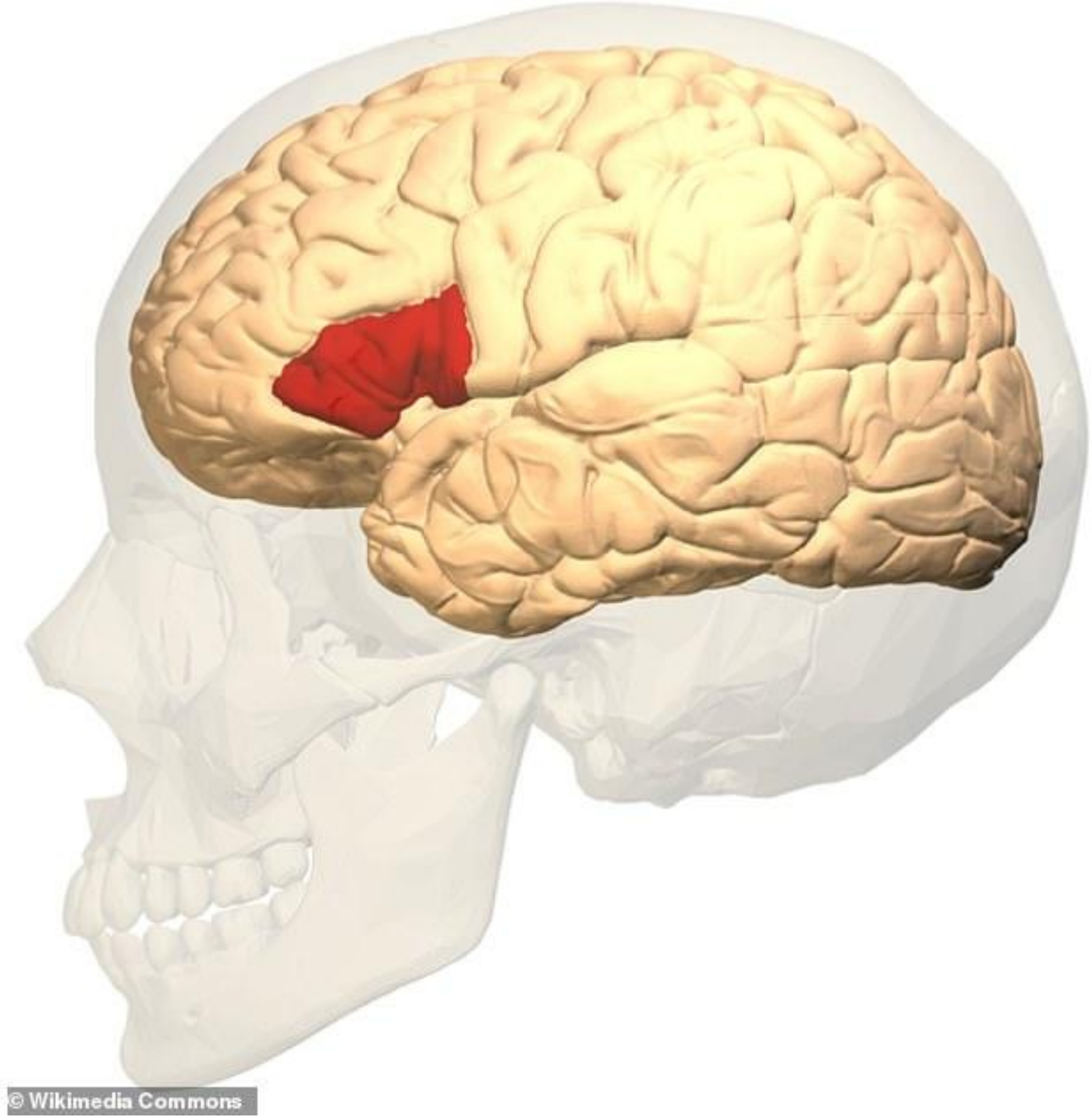
While Aristotle might have thought the brain existed only to cool the blood, serious investigation into how the brain produced thought began to gather speed in the 19th and 20th centuries.

One of the most important breakthroughs occurred in 1861 when a scientist called Paul Broca first visited a patient - a 21-year-old named Louis Victor Leborgne.

The patient suffered from a progressive loss of speech but no loss of the ability to understand or comprehend language.

When Leborgne died, Brocca removed his brain and found a lesion in the frontal lobe and hypothesized that this must be the site of language in the brain.

To this day, 'Broca's Area' is still associated with the ability to produce language.



Broca's Area, or Brodmann Areas 44 and 45, is still believed to play an important role in the production of language

What is a fMRI scan and how do they help us map the mind?

Functional Magnetic Resonance Imaging is a technique used to scan brains.

It works by detecting the small changes in blood blow that occur with brain activity.

Scanners use powerful magnetic fields and radiofrequency pulses to track the blood's movement.

Patients are awake and fully conscious during the procedure so they can conduct tasks while researchers observe changes in their brain.

While Broca's methods may seem simple, Professor Richard Bethlehem Director, of Neuroimaging at the Cambridge University Autism Research Centre, says modern methods are not all that different.

Professor Bethlehem told MailOnline: 'There are probably multiple ways that scientists interrogate the topology of brain function, very broadly you could say they often use either neuroimaging or lesion studies in combination with specific tasks designed to elicit specific functions.

'The premise of both branches is relatively simple though: find a task that you think elicits a specific function and then measure the activity/response or conduct the task with individuals that have specific lesions or brain damage in a region that might be crucial for that task.'

However, what has changed since Broca's day is that the neuroscientists' view of the brain is much less simple and views the brain's regions as far less isolated.

'Current consensus is more likely to suggest that there are often multiple brain regions involved in different functions and they may not even be selective to a singular function,' Professor Bethlehem added.

Some researchers think this theory should be pushed even further, and that we should reject the idea of there being particular brain regions for certain functions.

Dr Jeremy Skipper, principal investigator at UCL's Language, Action, and Brain Lab, told MailOnline that a function like language cannot be limited to just one part of the brain.

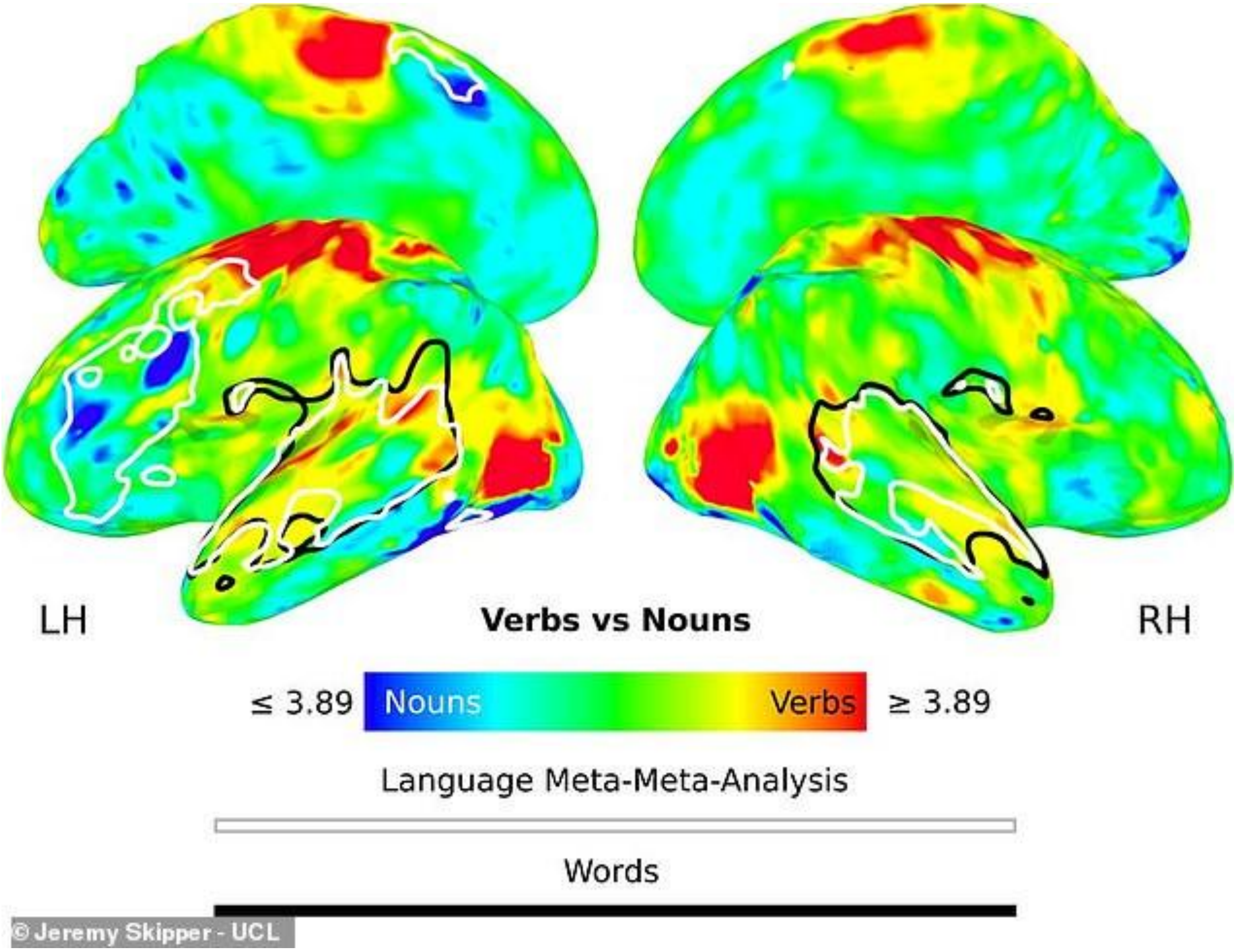
'What was classically called Broca's area,' Dr Skipper says, 'is not a real region but is composed of serval regions that have different cytoarchitectonic [cell structure] properties.

'That should be a clue from the very offset.'

To try and show that language was distributed throughout the brain, Dr Skipper and his team had subjects watch films while inside an fMRI scanner.

What he found was that 'when you listen to colour worlds, you activate colour regions. When you listen to words that involve motion, you activate motor processing regions.'

Dr Skipper believes that the regions typically called 'language regions' are really just processing connectivity hubs which are all that are left behind when scientists average out lots of individual brain scans.



Dr Skipper's brain scans show that nouns (blue) and verbs (red) tend to activate very different regions of the brain than the typical language processing areas (outlined in white)

READ MORE: [Brain images just got 64 million times sharper thanks to new turbocharged MRI scans]

For example, imagine a city that still used telephone switchboards. If you averaged out all phone activity all that would be left behind would be the call centres which coordinate everyone's phones.

If that was all the data left, you might think that phone calls only happen inside switchboards but then you would miss out on understanding who is talking to who.

The radical conclusion, Dr Skipper says, is that 'there is no speech region in the brain.'

'There is no language comprehension region,' he added.

'Brain regions on their own are not functional modules, their functions are determined by what other brain regions they are talking to.'

Like all scientific theories, Dr Skipper's theory isn't without its critics and some researchers still believe that there are brain functions that can be restricted to specific areas.

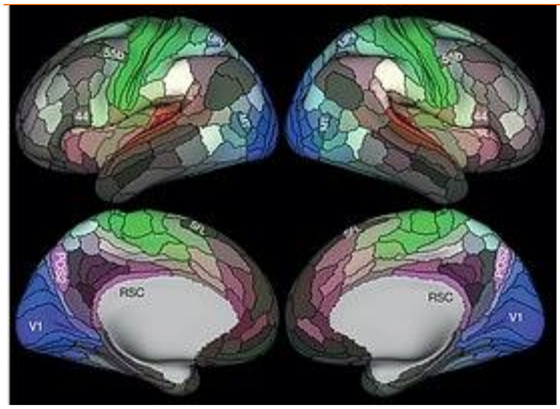
Some point to lesion studies and cases like strokes where specific brain damage does cause specific cognitive issues with associated functions.

However, theories like these are gaining traction, and there is a growing assumption that functions of the brain are far more spread out than Broca would have ever considered.

Professor Bethlehem added: 'I think nowadays we tend to think of the brain as a distributed network where there are certainly hubs and areas that have preference for certain functions, but they are not isolated chunks of information processing tissue.'

Share or comment on this article: Delving inside the mind: Incredible graphics

70,000 thoughts processed every day



These incredible graphics show in detail how each part of the brain contributes to the functions and abilities that make us all uniquely human. However, not all scientists think it can be that simple.

Dr. Kinney, Dr. Edlow and their team then double-checked this work against so-called 'in vivo 7-tesla resting-state functional MRI' data, which had been collected safely from living patients who had volunteered for **the Human Connectome Project**.

Based on their mapping work, as they wrote in their new publication in the journal **Science Translational Medicine**, the researchers believe that **the ventral tegmental area, or VTA, in the midbrain** will be a key region for treatments to help with wakefulness.

The VTA is known to be the origin of a number of nerve fibers that produce dopamine, the hormone involved in 'rewarding' the brain and happiness.

'Our connectivity results suggest that stimulation of the ventral tegmental area's dopaminergic pathways has the potential to help patients recover from coma,' Dr. Edlow said in a statement.

'This hub node is connected to many regions of the brain that are critical to consciousness,' he pointed out.

Dr. Edlow, an associate director of the Center for Neurotechnology and Neurorecovery (CNTR) at Boston's Mass General and an associate professor of Neurology at Harvard Medical School, said he hopes to continue providing more detailed neural maps in an effort to help patients who have suffered from a loss of consciousness.

'We envision that these connectivity maps will allow us to piece together, one individual at a time, the combination of connections that are necessary and sufficient to recover consciousness,' he said.

Researchers Find the Connection Between Theta Connectivity and Intelligence in a Brand New Study

Neural network and robot. (Representative Cover Image Source: Getty Images | Yuichiro Chino)© Front Page Detectives

Researchers Find the Connection Between Theta Connectivity and Intelligence in a Brand New Study

If a human wants to become intelligent, they don't need to read books or understand theories; they just need to have good coordination in their head. That's what a new study published in the Journal of Experimental Psychology claims. Researchers conducted a set of tests on some volunteers, which revealed that the stronger an individual's theta connectivity, the better their chances of being intelligent. Different regions start coordinating when the brain performs tasks that need more than normal cognitive skills. The syncing of these signals is called theta connectivity. Psychologist Anna-Lena Schubert of Johannes Gutenberg University Mainz in Germany claims that people who displayed better proficiency in these cognitive skills had more synchronized signals in their brains compared to those with less proficiency, according to Science Alert.

Active human brain - stock photo (Representative Image Source: Getty Images | Photo by PM Images)© Front Page Detectives

Objective of the Study

Researchers recruited 148 people between the ages of 18 and 60 and made them solve three tests, focusing on their memory and intelligence. The study's objective was to figure out whether theta connectivity has any relation with cognitive control. Cognitive control refers to the ability of the human brain to adjust and adapt its thoughts, behavior, and emotions to different situations. Experts wanted to see whether the brain signals coordinated with each other when an individual faced different tasks requiring distinct skills. These signals are a series of slow waves in the band of 4 to 8 Hertz called theta signals.

Researchers recording participants' brain activity via a non-invasive electroencephalography (EEG) cap. (Representative Image Source: Getty Images | romaset)© Front Page Detectives

Tests Given to Participants

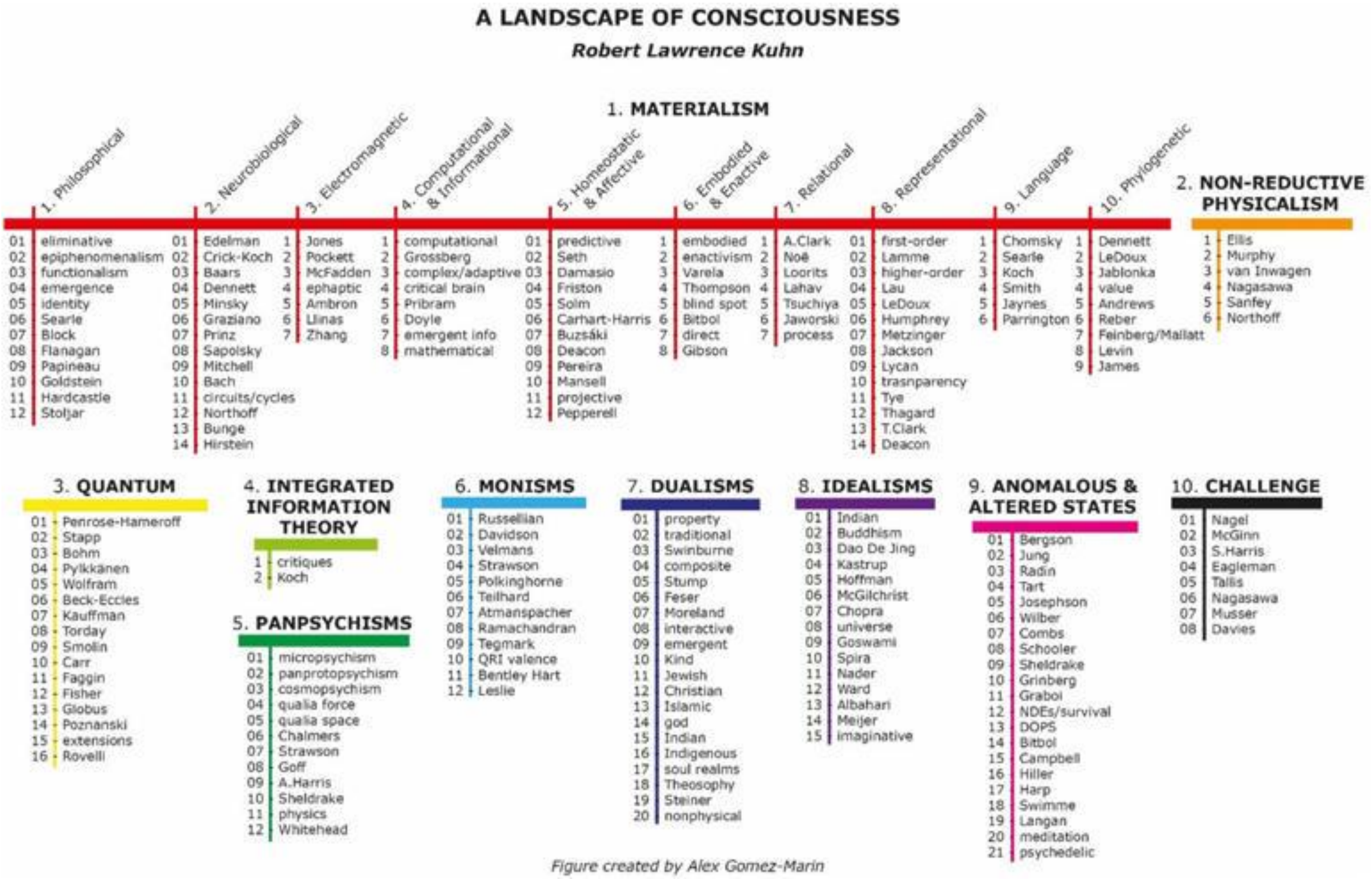
Researchers gave each participant a non-invasive electroencephalography (EEG) cap to record participants' brain activity. For the first task, they provided the participants with numbers from one to nine and asked whether they were below five or more than five. In the second one, the participants were given Navon figures and had to identify the big and small shapes in them. For the third one, researchers gave participants many pairs of numbers and letters, where they had to answer whether the alphabet was a vowel or a consonant, or if the number was more or less than five. Researchers wanted to observe the impact of changing rulesets in the brains of participants.

Results of the Tests

Experts observed brain signals coordinating with each other every time they changed the tasks. It proved that theta connectivity was indeed linked to cognitive control. They also noted that the syncing was stronger in people who performed better in these intelligence and memory tests. "People with stronger midfrontal theta connectivity are often better at maintaining focus and tuning out distractions, be it that your phone buzzes while you're working or that you intend to read a book in a busy train station," Schubert added. "We did not expect the relationship to be this clear."

Researchers noted that the coordination wasn't as strong during cue recognition as it was during actual responses. The syncing was weaker when individuals were being instructed on different rules or were preparing for the next task with distinct requirements. They were strong when the individual was solving different problems. It implies that intelligence may not be dependent on things one does before the task, but on the ability to handle things on the spot. "These insights significantly advance theoretical models of intelligence, highlighting the critical role of specific aspects of cognitive control in cognitive abilities," the researchers wrote. Experts hope to utilize these findings to learn further about the relationship between brain signal coordination and cognition. They believe that these insights could also lead to new diagnostic techniques for ailments related to the brain.

A landscape of consciousness: Neurophysiologist presents diverse theories and taxonomy of proposed solutions



A taxonomy of consciousness explanations. Credit: Robert Lawrence Kuhn (2024). Created by Robert Lawrence Kuhn and Alex Gomez-Marin. A landscape of consciousness. Note: Categories 1–10 in the Figures correspond to sections 9-18 in the text. To convert from categories/theories in the Figures to sections/theories in the text, add eight (+8). Conversely, to convert from sections/theories in the text to categories/theories in the Figures, subtract eight (−8). Progress in Biophysics and Molecular Biology (2024). DOI: 10.1016/j.pbiomolbio.2023.12.003

"Out of meat, how do you get thought? That's the grandest question." So said philosopher Patricia Churchland to Robert Lawrence Kuhn, the producer and host of the acclaimed PBS program, Closer to Truth.

Now Kuhn, a member of FQxI's scientific advisory council, has published a taxonomy of proposed solutions to, and theories regarding, the hard problem of consciousness. He produced the organizing framework in order to explore their impact on meaning, purpose and value (if any), AI consciousness, virtual immortality, survival beyond death, and free will. Kuhn's "A landscape of consciousness: Toward a taxonomy of explanations and implications" was published in the journal *Progress in Biophysics and Molecular Biology* in its August 2024 issue.

FQxI and Closer To Truth's Robert Lawrence Kuhn (middle) interviews cosmologists Alan Guth (left) and Andrei Linde (right) during an FQxI meeting. Credit: Robert Lawrence Kuhn (2009).

Closer To Truth has aired 333 television episodes since 2000, including 30 in partnership with the Foundational Questions Institute, FQxI. Kuhn's article is the product of some of his many in-depth interviews with experts over the decades.

"I have discussed consciousness with over 200 scientists and philosophers," says Kuhn, who is himself trained in neurophysiology. "Landscape is the product of a lifetime."

The article begins with the classic mind-body problem: How do the felt experiences in our minds relate to the neural processes in our brains? How do mental states, whether sensory, cognitive, emotional, or even noumenal (self-less) awareness, correlate with brain states?

"Although there are families of mind-body problems, I focus tightly on phenomenal consciousness: our inner awareness, 'what it feels like to be' something," says Kuhn.

Kuhn presents diverse theories of consciousness from materialist/physicalist to nonmaterialist/nonphysicalist. These are categorized as: Materialism Theories (philosophical, neurobiological, electromagnetic field, computational and informational, homeostatic and affective, embodied and enactive, relational, representational, language, phylogenetic evolution); Non-Reductive Physicalism; Quantum Theories; Integrated Information Theory; Panpsychisms; Monisms; Dualisms; Idealisms; Anomalous and Altered States Theories; and Challenge Theories.

Seeking insights

"My purpose must be humble: collect and categorize, not assess and adjudicate," says Kuhn. "Seek insights, not answers." Kuhn produced the organizing framework for these diverse theories of consciousness in order to explore their impact on "ultimate questions," such as meaning, purpose and value (if any), AI consciousness, virtual immortality, survival beyond death, and free will, he says.

"Understanding consciousness at this point cannot be limited to selected ways of thinking or knowing, but should seek expansive yet rational diversity."

Having produced an article of around 175,000 words, Kuhn found that his opinions on certain proposals had evolved. "My own hunch, right here, right now might be something of a Dualism-Idealism mashup. Second place might go to some form of Quantum Consciousness, triggered by writing this paper and surprising me. Third place, counterintuitively, to a kind of Eliminative Materialism/Illusionism, combined with Neurobiological and Representational Theories."

But, adds Kuhn, "Smart, serious folks believe radically different theories; what I believe doesn't much matter."

More information: Robert Lawrence Kuhn, A landscape of consciousness: Toward a taxonomy of explanations and implications, *Progress in Biophysics and Molecular Biology* (2024). DOI: 10.1016/j.pbiomolbio.2023.12.003

You can read more about Kuhn's writing process and his thoughts on consciousness in his FQxI article: "A Landscape of Consciousness."

Provided by Foundational Questions Institute, FQXi

A Landscape of Consciousness

OCTOBER 24, 2024

by Robert Lawrence Kuhn

"Out of meat, how do you get thought? That's the grandest question."

So philosopher Patricia Churchland once told me, when speaking about the hard problem of consciousness.

Recently, I have applied FQxI ways of thinking to my original field of neuroscience and consciousness studies. In August 2024, I published a comprehensive review of theories of consciousness in *Progress in Biophysics and Molecular Biology* ("A Landscape of Consciousness: Toward a Taxonomy of Explanations and Implications" Volume 190, August 2024, Pages 28-169).

Landscape is the product of a lifetime. To give you my background: My PhD was in neurophysiology—thalamocortical evoked potentials—at UCLA's Brain Research Institute, 1968. I created, executive produce, write and present *Closer to Truth* (CTT); broadcast on PBS stations nationally, CTT was launched in 2000 and re-launched in 2008 (co-created, produced and directed by Peter Getzels). CTT tracks my global journey to discover state-of-the-art ideas about raw existence and human sentience. In recent years, CTT has become a global resource via the *Closer To Truth* website and *Closer To Truth* YouTube channel.

I was an early FQxI member and am on FQxI's Scientific Advisory Council. Since 2007, FQxI has been a special content partner of CTT on Cosmos (cosmology/fundamental physics, philosophy of science), Life (philosophy of biology), Consciousness (brain/mind, philosophy of mind), and Meaning (global philosophy of religion, critical thinking). Together, we have developed and promoted state-of-the-art ideas and over-the-horizon thinking: cosmology (Iceland, 2007), time (Norway/Denmark, 2011), information (Vieques, 2014), and physics of the observer/what happens (Banff, 2016). We have broadcast about 30 CTT FQxI-derived TV episodes and streamed over 500 CTT FQxI-derived videos. I have been privileged to bring FQxI members, topics, ways of thinking to broad international audiences.

Moreover, roughly one-third of *Closer To Truth* focuses on mind/brain topics; I have discussed consciousness with over 200 scientists and philosophers on the *Closer To Truth* website and *Closer To*

Truth YouTube Channel. In the photo above, I am interviewing FQxI President Anthony Aguirre. In the photo below, you see me with FQxI cosmologists Alan Guth (left) and Andrei Linde (right).

This is how I begin Landscape:

"Explanations of consciousness abound and the radical diversity of theories is telling. Explanations, or theories, are said to work at astonishingly divergent orders of magnitude and putative realms of reality."

In Landscape, I present diverse theories of consciousness from materialist/physicalist to nonmaterialist/nonphysicalist. Categories: Materialism Theories (philosophical, neurobiological, electromagnetic field, computational and informational, homeostatic and affective, embodied and enactive, relational, representational, language, phylogenetic evolution); Non-Reductive Physicalism; Quantum Theories; Integrated Information Theory; Panpsychisms; Monisms; Dualisms; Idealisms; Anomalous and Altered States Theories; Challenge Theories. There are many subcategories, especially for Materialism Theories. (See Figure below, produced by Alex Gomez-Marin and myself.) Each explanation is self-described by its adherents, critique is minimal and only for clarification, and there is no attempt to adjudicate among theories.

I seek an organizing framework for these diverse theories of consciousness and to explore their impact on "ultimate questions." I have two central theses: (i) understanding consciousness at this point cannot be limited to selected ways of thinking or knowing, but should seek expansive yet rational diversity, and (ii) ultimate questions related to consciousness, such as meaning/purpose/value (if any), AI consciousness, virtual immortality, survival beyond death, free will, and the like., cannot be understood except in the light of particular theories of consciousness.

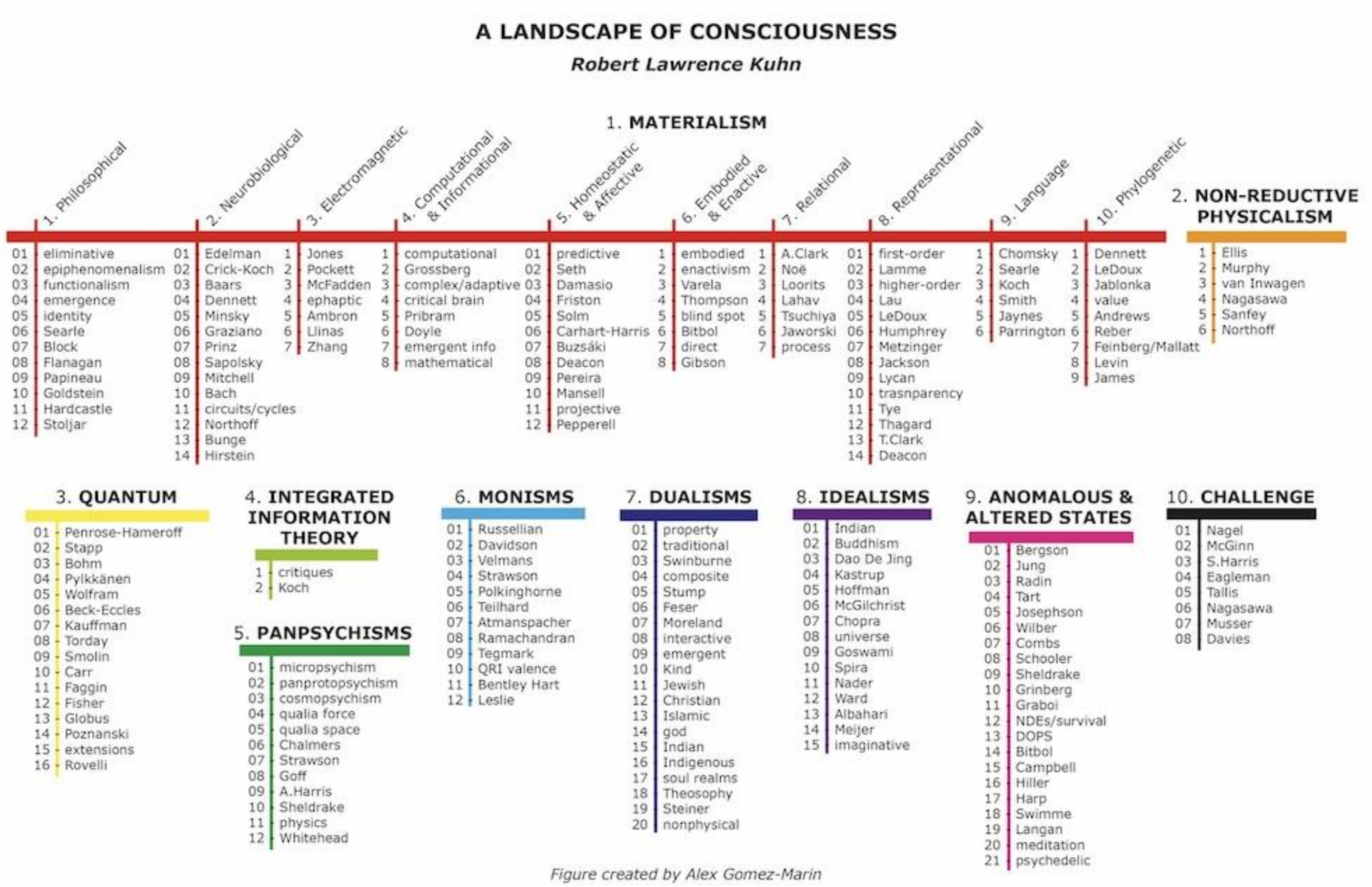


Figure created by Alex Gomez-Marin

Here's the backstory. The journal, *Progress in Biophysics and Molecular Biology* (PBMB), invited me to write an article on consciousness broadly. (I was reluctant, imagining what it would entail—even so, I wildly underestimated.) In my first draft, I included, along with all the hard science, philosophical and theological explanations: complex philosophical theories (e.g., non-reductive physicalism, emergent dualism, cosmopsychism) and diverse religious theories (i.e., Hindu, Buddhist, Daoist, Judaic, Christian, Islamic, Indigenous). The anonymous peer reviewer, while generously appreciative of the scientific theories (and offering helpful suggestions), recommended cutting the purely philosophical and theological theories—which, he said, journal readers would not much care about. I responded by agreeing that journal readers would likely not care, but that I surely did care, and because I was going to do this Landscape once in my life, I must make it as complete as I, in my idiosyncratic way, saw fit. If PBMB did not want to publish it with all the theories, I'd totally understand and seek another venue. To PBMB's credit, they agreed to publish it all, nothing cut. I note in Landscape that my purpose must be humble: collect and categorize, not assess and adjudicate. Seek insights, not answers. Unrealistically, I'd like to get all the theories, at least all contemporary theories that are sufficiently distinct with explanations that can surmount an arbitrary hurdle of rationality or conceivability. Falsification or verification is not on the agenda. I'm less concerned about the ontological truth of explanations/theories than in identifying them and then locating them on a "Landscape" ("A" Landscape, not "The" Landscape) to enable categorization and assess relationships. Thus, Landscape is not about how consciousness is measured or evolved or even how it works, but about what consciousness *is* and what difference it makes.

It starts as the classic "mind-body problem:" How do the felt experiences in our minds relate to the neural processes in our brains? How do mental states, whether sensory, cognitive, emotional, or even noumenal (self-less) awareness, correlate with brain states? Although there are families of mind-body problems, I focus tightly on phenomenal consciousness: our inner awareness, "what it feels like to be" something. I avoid areas adjacent to phenomenal consciousness: perception, cognition, emotion (per se), subconscious, self, intelligence, decision making, etc.—each of which is a philosophical-scientific matrix of its own. Because Landscape is structured by theories of consciousness, not by philosophical questions, each theory sets its own agenda for dealing with phenomenal consciousness.

The Landscape explanations or theories that I want to draw are as broad as possible, including those that cannot be subsumed by, and possibly not even accessed by, the scientific method. This freedom from constraint, as it were, is no excuse for wooly thinking. Standards of rationality and clarity of argument must be maintained even more tenaciously, and bases of beliefs must be specified even more clearly.

Theories overlap; some work together. Moreover, while a real-world "landscape of consciousness," even simplified, would be drawn with three dimensions (at least), with multiple kinds and levels of nestings—a combinatorial explosion (and likely no closer to truth)—I satisfice with a one-dimensional toy-model. I array all the theories on a linear spectrum, simplistically and roughly, from the "most physical" on the left (at the beginning) to the "least physical" on the right (near the end). (I have two final categories after this spectrum.) The Figure above provides an overall outline of Landscape.

The physicalism assumed in Materialism Theories of consciousness is characterized by naturalistic, science-based perspectives, while non-materialism theories have various degrees of nonphysical perspectives outside the ambit of current science, and, as noted, in some cases not subject to the scientific method of experimentation and replicability.

Please do not assume that in Landscape the relative importance of a theory is proportional to the relative size of its description. Shorter can be stronger. For each description I feel the tension between conciseness and completeness. Furthermore, several are not complete theories in themselves but ways to think about consciousness that strike me as original and perhaps insightful. I appreciate the many readers who have written in response to Landscape, seeing impact in the field. In his review, Alex Gomez-Marin writes: "When was the last time you read a piece cordially inviting philosophy, neuroscience, quantum physics, psychical research, theology, and religion to the same table?" Unexpected is Stuart Kauffman's: "By publishing the Landscape you shall have changed it." While I don't know if that's true, I do know that several observe—and not with pleasure—the proliferation, not the pruning, of theories. I've said about theories of consciousness something akin to what I've said about varieties of religion (with tongue only partially in cheek): "It's not that we have too many; it's that we have one too few."

In his influential paper, "What Is It Like to Be a Bat?," Thomas Nagel offers wisdom: "Without consciousness the mind-body problem would be much less interesting. With consciousness it seems hopeless."

"Hopeless," to me, is invigorating; I'm up for an FQxI-style "hopeless challenge." Please take "A Landscape of Consciousness" as my personal journey of consciousness; idiosyncratic, to be sure; not all for everyone, not set in cement.

So, after more than five decades of thinking, hundreds of *Closer To Truth* interviews and discussion, and a consuming ~175,000 words in Landscape, did anything surprise me? Prior, I had utterly dismissed quantum theories of consciousness and utterly rejected that psychedelics could provide veridical perspectives of ontological reality. As I note, writing Landscape has put a hairline fracture in my utterly bone-strength worldview.

What do I personally believe? While I am often asked this question, and I appreciate why, I still do not like it. Smart, serious folks believe radically different theories; what I believe doesn't much matter. In my Landscape review, I did not want to defend, or even to offer, my own view because it might skew perceptions of the entire enterprise. I try to present each theory as accurately and persuasively as I can, usually with the words of its creator. In writing each theory, I tried to "inhabit" it, to imagine it was my own theory, which I wanted to world to appreciate. However, I came to recognize that if I did not say something about what I thought, my not-so-veiled omission could also seem like a kind of "hidden agenda." So, right at the very end of Landscape, I have this short paragraph:

"Me, I just don't know...My own hunch, right here, right now—if I'm coerced to disclose it and for what little it's worth—might be something of a Dualism-Idealism mashup. (I can describe; I dare not defend.)"

with this footnote:

"Second place might go to some form of Quantum Consciousness, triggered by writing this paper and surprising me. Third place, counterintuitively, to a kind of Eliminative Materialism/Illusionism, combined with Neurobiological and Representational Theories."

Thirty-nine words in the text, 33 in the footnote; 72 words all together—out of ~175,000. I've come to love the blizzard-like storm of theories, luxuriating in it—but only because I love consciousness and see it as the central question of existence and sentience (no matter what the ultimate answer), which is why I do not want to limit its meaning or mechanism (at this moment).

Finally, as we witness on *Closer To Truth*, I'm gratified that a passion for consciousness can bring together people from different nations, regions, religions, races, ethnicities, genders, educational levels, income levels. We human beings are all united by ultimate questions.

****

Robert Lawrence Kuhn is creator, executive producer, writer and host of *Closer To Truth* on science and philosophy (Cosmos. Life. Mind. Meaning.).

Neuroscientists identify a shared brain circuit for creativity

[Adobe Stock]© PsyPost

New research published in *JAMA Network Open* has identified a common brain circuit linked to creativity across multiple domains, such as music, writing, drawing, and idea generation. The findings indicate that damage to this circuit—caused by brain injuries or neurodegenerative diseases—may sometimes enhance creative abilities. This counterintuitive result adds to growing evidence that creativity is not only supported by specific brain networks but might also be released when certain regions are disrupted.

“We wanted to answer the questions, ‘What brain regions are key for human creativity and how does this relate to the effects of brain injuries?’” said study author Isaiah Kletenik, an assistant professor at Harvard Medical School, researcher at the Center for Brain Circuit Therapeutics, and associate neurologist at Brigham and Women's Hospital.

"As a cognitive neurologist, I take care of patients who have brain injuries or dementia which you would assume should only cause dysfunction but there are these rare situations in which people have a progressive brain disease and have new onset creativity. The first author, Julian Kutsche has a background in both neuroscience and in the study of music so this topic was a perfect combination of his research interests."

To address this question, Kletenik and his colleagues at the Center for Brain Circuit Therapeutics at Brigham and Women's Hospital examined data from 857 participants who had taken part in 36 brain imaging studies of creativity. Each of these studies used functional MRI to identify brain regions activated during creative tasks compared to control tasks. The researchers then applied a technique known as coordinate network mapping to analyze the functional connectivity of those brain areas using resting-state data from 1,000 healthy adults.

Although the specific brain regions activated during creativity tasks varied widely across studies, the researchers found that a large majority—86%—shared a connection to a common brain circuit. This circuit was characterized by negative functional connectivity with a particular region in the brain known as the right frontal pole. That is, when creative brain regions were active, the right frontal pole tended to be less active, suggesting a possible suppressive or regulatory role in creative thinking.

This finding was consistent across different types of creative tasks, different statistical thresholds, and different levels of data analysis, making it a robust result. To ensure their findings were not due to random chance or general brain activity, the researchers also compared their data with brain networks involved in unrelated tasks like working memory. The creativity circuit stood out as distinct.

"Many complex human behaviors such as creativity don't map to a specific brain region but do map to specific brain circuits," Kletenik told PsyPost. "Some people with neurologic diseases experience a new onset of creative behavior and show specific patterns of damage that align with a brain circuit associated with creativity."

"We found that many complex human behaviors such as creativity don't map to a specific brain region but do map to specific brain circuits," added co-senior author Michael D. Fox, who founded and leads the Center for Brain Circuit Therapeutics, and helped develop the techniques of coordinate and lesion network mapping employed in this work.

To validate their results, the team looked at three independent lines of evidence: brain activation patterns in an additional 30 fMRI studies, data from patients with focal brain lesions, and data from patients with neurodegenerative diseases. In each case, the results supported the existence of a common creativity circuit.

One of the more intriguing validations came from examining patients with brain lesions in different parts of the frontal lobe. Patients with lesions in areas negatively connected to the right frontal pole—such as the lateral frontal lobe—tended to show reduced creativity. But those with lesions in the right frontal pole itself sometimes showed enhanced creative output, suggesting that damage to this area may reduce self-monitoring or internal censorship, allowing freer expression of creative ideas.

Similarly, the researchers found that certain neurodegenerative diseases—such as semantic variant primary progressive aphasia (svPPA) and the behavioral variant of frontotemporal dementia (bvFTD)—are associated with increased creativity. The brain atrophy patterns in these diseases aligned strongly with the identified creativity circuit, especially in regions negatively linked to the right frontal pole. In other conditions, such as Parkinson's disease or nonfluent aphasia, where increased creativity is less commonly observed, the alignment was weaker.

The team also analyzed data from patients with frontotemporal dementia who had experienced a surprising boost in artistic abilities. These individuals showed brain atrophy in areas that overlapped with the newly defined creativity circuit, adding further evidence that disruption of this network may allow creativity to emerge in unexpected ways.

Taken together, these results suggest that the right frontal pole may act as a kind of gatekeeper, suppressing spontaneous or unconventional thinking in favor of rule-based decision-making. When that control is reduced—whether by task-induced deactivation, injury, or degeneration—creative behavior may become more prominent. This idea is consistent with previous studies that found lower activity in parts of the frontal cortex during improvisation and other free-form tasks.

“The most surprising finding was that all the different forms of creative tasks shared a specific regional reduced activity on brain imaging,” Kletenik explained. “In functional neuroimaging we often focus on areas where there’s increased activation, but in this study the key finding was that all the creative tasks seem to share reduced activation in the right frontal pole. This part of your brain is important for monitoring and rule-based behaviors. Creativity may depend on inhibiting self-censoring assessments that could then allow free association and idea generation to flow more freely.”

While the findings are compelling, the researchers stress that creativity is complex and likely involves many different brain systems. Their work focused on identifying a shared neural pathway across studies, not on capturing every possible variation. Moreover, the observed relationships are correlational; while lesion data suggest a causal role for the right frontal pole, more direct interventions—such as brain stimulation studies—are needed to confirm this.

“These findings rely on retrospective analyses and do not represent the entire neural circuitry involved in creativity – there are many different parts of the brain necessary for completing different creative tasks,” Kletenik noted.

Still, the findings open the door to new ways of thinking about the brain and its relationship to creativity. They may also help explain why some people report an increase in artistic or musical output after brain injuries or during early stages of neurological disease. This phenomenon, known as paradoxical functional facilitation, suggests that impairments in one area of the brain can sometimes lead to gains in another.

“These findings could help explain how some neurodegenerative diseases might lead to a paradoxical increase in creativity,” Kletenik explained. “We are learning more about neurodiversity and how brain changes that are considered pathological may improve function in some ways

The study, “Mapping Neuroimaging Findings of Creativity and Brain Disease Onto a Common Brain Circuit,” was authored by Julian Kutsche, Joseph J. Taylor, Michael G. Erkkinen, Haya Akkad, Sanaz Khosravani, William Drew, Anna Abraham, Derek V. M. Ott, Juliana Wall, Alexander Li Cohen, Andreas Horn, Wolf-Julian Neumann, Isaiah Kletenik, and Michael D. Fox.

Mapping Neuroimaging Findings of Creativity and Brain Disease Onto a Common Brain Circuit

Julian Kutsche, MA[1,2,3]; Joseph J. Taylor, MD, PhD[2,4]; Michael G. Erkkinen, MD[3]; et al

Key Points

Question Does creativity map to a specific brain circuit, and does damage to this circuit align with creativity changes that occur in brain disease?

Findings This study using network mapping of meta-analytic data involving 857 participants found that brain regions activated by creativity tasks mapped to a human brain circuit centered on the right frontal pole. Damage to this circuit aligned with both decreases and paradoxical increases in creativity observed across multiple different brain diseases.

Meaning Findings from this study suggest that creativity maps to a specific brain circuit in healthy individuals and that damage to this circuit in individuals with certain neurodegenerative diseases may explain their paradoxical increases in creativity; further research is needed to confirm and validate these findings.

Abstract

Importance Creativity is important for problem solving, adaptation to a changing environment, and innovation. Neuroimaging studies seeking to map creativity have yielded conflicting results, and studies of patients with brain disease have reported both decreases and paradoxical increases in creativity, leaving the neural basis of creativity unclear.

Objective To investigate the brain circuit underlying creativity and assess its association with brain injury and neurodegenerative disease.

Design, Setting, and Participants This study examined neuroimaging coordinates from a meta-analysis of 36 studies published between 2004 and 2019 associated with increased activity during creative tasks in healthy participants. A validated method termed *coordinate network mapping* and a database of resting-state functional connectivity from 1000 healthy individuals were used to test whether these coordinates mapped to a common brain circuit. Specificity was assessed through comparison to random coordinates and coordinates from working memory tasks in healthy participants. Reproducibility was assessed using an independent dataset of coordinates from additional studies of creativity in healthy participants. Finally, alignment with effects of focal brain damage on creativity was tested using data from patients with brain lesions and coordinates of brain atrophy from 7 different neurodegenerative disorders.

Main Outcomes and Measures The primary outcomes were creativity or no creativity and alignment with a creativity circuit or no alignment.

Results Creativity tasks activated heterogenous locations, with coordinates scattered across many different brain regions (415 coordinates derived from 857 healthy participants; pooled mean [SD] age, 24.1 [6.91] years; 461 [54%] female). However, these activation coordinates were part of a common brain circuit, defined by negative connectivity to the right frontal pole. This result was consistent across creative domains, reproducible in an independent dataset (383 coordinates derived from 691 participants) and specific to creativity when compared with random gray matter coordinates (n = 415) or coordinates activated by working memory tasks (3072 coordinates derived from 2900 healthy participants). Damage to this creativity circuit by lesions (n = 56 patients) or neurodegenerative disease (2262 coordinates derived from 4804 patients) aligned with both decreases and increases in creativity observed in these disorders.

Conclusions and Relevance Findings from this study suggest that brain regions activated by creativity tasks map to a brain circuit defined by negative functional connectivity to the right frontal pole. Damage to this circuit aligned with changes in creativity observed in individuals with certain brain diseases, including paradoxical creativity increases.

Introduction

Creativity involves generation of novel and useful products or ideas in a changing environment.[1] It is important for problem solving and is the foundation for technological, artistic, and cultural innovation. The neuroanatomical substrate of creativity has been a topic of long-standing neuroscientific interest[2] but may also be relevant for medical practitioners in understanding symptoms among patients with brain disease.[3] Some evidence suggests creativity can be affected by brain disease, including focal brain damage[4-6] and neurodegenerative disorders, as well as by treatments for brain disease.[7-10] Brain disease can lead to decreases or paradoxical increases in creativity.[8,11-16] Despite long-standing interest in creativity and its relationship to brain disease,[17] it remains unclear which brain regions are most important for creativity.

The neural underpinnings of creativity have been studied in 2 main ways: (1) neuroimaging studies in healthy participants and (2) alterations in creativity in patients with brain disease. Both approaches have identified brain regions involved in creativity, but results have been heterogenous across studies. Among healthy participants, different creative tasks tend to activate different brain regions, suggesting that different types of creativity may map to different neuroanatomy.[18-21] However, these results do not preclude the possibility of a domain-general neuroanatomical substrate for creativity, the existence of which remains an ongoing topic of debate.[22,23] Other neuroimaging studies have examined the relationship between creativity and brain network connectivity.[24-30] These articles suggest that creativity may map better to brain networks than individual brain regions, but exactly which network and how these networks relate to brain regions activated by creativity tasks remains unclear.

Among patients, different brain regions have been implicated across different disorders.[4,31-33] For example, creativity increases have been observed in patients with semantic variant primary progressive aphasia (svPPA),[12] which is associated with atrophy to the anterior lateral temporal lobes. However, associations of this brain region with creativity are not consistent,[34] and other brain regions, such as the dorsolateral prefrontal cortex, have been associated with facilitation of creativity in patients with neurodegenerative disease.[35,36]

The aim of this study was to take a network approach to localizing creativity in the brain across different creative domains, including music, writing, drawing, and divergent thinking. This network mapping approach can test whether heterogenous results observed across different studies map to a connected brain circuit.[37] The approach has been validated for use with neuroimaging coordinates activated by different tasks,[38,39] lesion locations,[40-43] and coordinates of brain atrophy associated with neurodegeneration.[44] It thus has the potential to identify convergent, domain-general neurocircuitry for creativity.

Methods

This study using network mapping of meta-analytic data followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline. A full description of methods is given in the eMethods in Supplement 1. In brief, we extracted brain coordinates activated by creativity tasks (vs control tasks) from 36 prior functional magnetic resonance imaging (fMRI) studies published between 2004 and 2019 that were included in a meta-analysis.[20]

Statistical Analysis

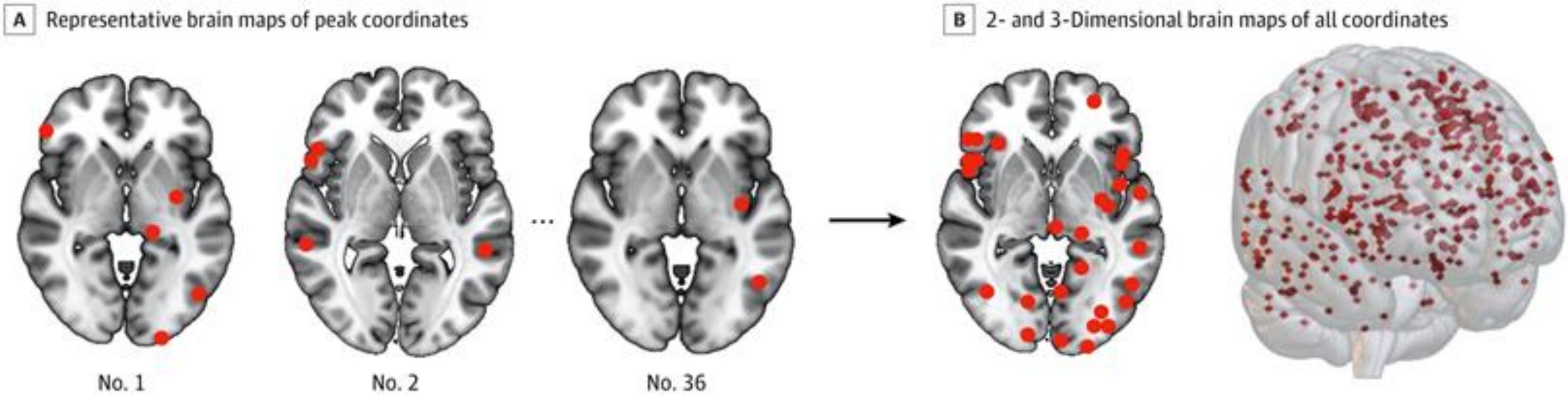
A validated method termed *coordinate network mapping* and resting-state functional connectivity data from 1000 heathy participants were used to identify the brain circuit functionally connected to each set of coordinates, resulting in 36 network connectivity maps.[38,39,45] To identify brain regions common to these 36 maps, we thresholded the network maps ($t \geq 5$), binarized them, then overlaid them. To identify brain regions that were statistically consistent across these 36 maps, we performed a 2-tailed 1-sample t test and identified significant clusters after correction for multiple comparisons (family-wise error rate, $P < .05$). To identify brain regions that were specific to these 36 network maps, we performed a 2-sample t test comparing the 36 maps derived from studies of creativity to (1) 36 network maps derived from random gray matter coordinates and (2) 150 network maps derived from fMRI studies of working memory. We again identified significant clusters after correction for multiple comparisons (false discovery rate $P < .05$). We used a conjunction analysis of the above results to identify any brain network connections that were both sensitive and specific for creativity. We validated the results by comparing them to coordinates activated by creativity tasks in an independent set of fMRI experiments (n = 30 studies) (eTable 3 in Supplement 1),[18,19] the location of brain lesions associated with changes in performance on creativity tasks (n = 56 patients with lesions),[31] and the locations of brain atrophy in patients with neurodegenerative disease associated with changes in creativity (n = 189 studies[46-51]) (eTable 5 in Supplement 1).

The functional connectivity data equivalent to those used in this study are available online through the Harvard Dataverse.[52] The fMRI and atrophy data used in this study are publicly available and obtained from published medical literature (eTables 1, 3, 4 and 5 in Supplement 1). The threshold for statistical significance was a 2-tailed $P < .05$. Statistical neuroimaging analyses were performed in Matlab, version 2022b (Mathworks Inc) and using FSL software (version 5.0.10). Brain images were created using MRicroGL (version 12.7.6), FSLeyes (version 1.4.6), and Surf Ice (version 12.7.6) software.

Results

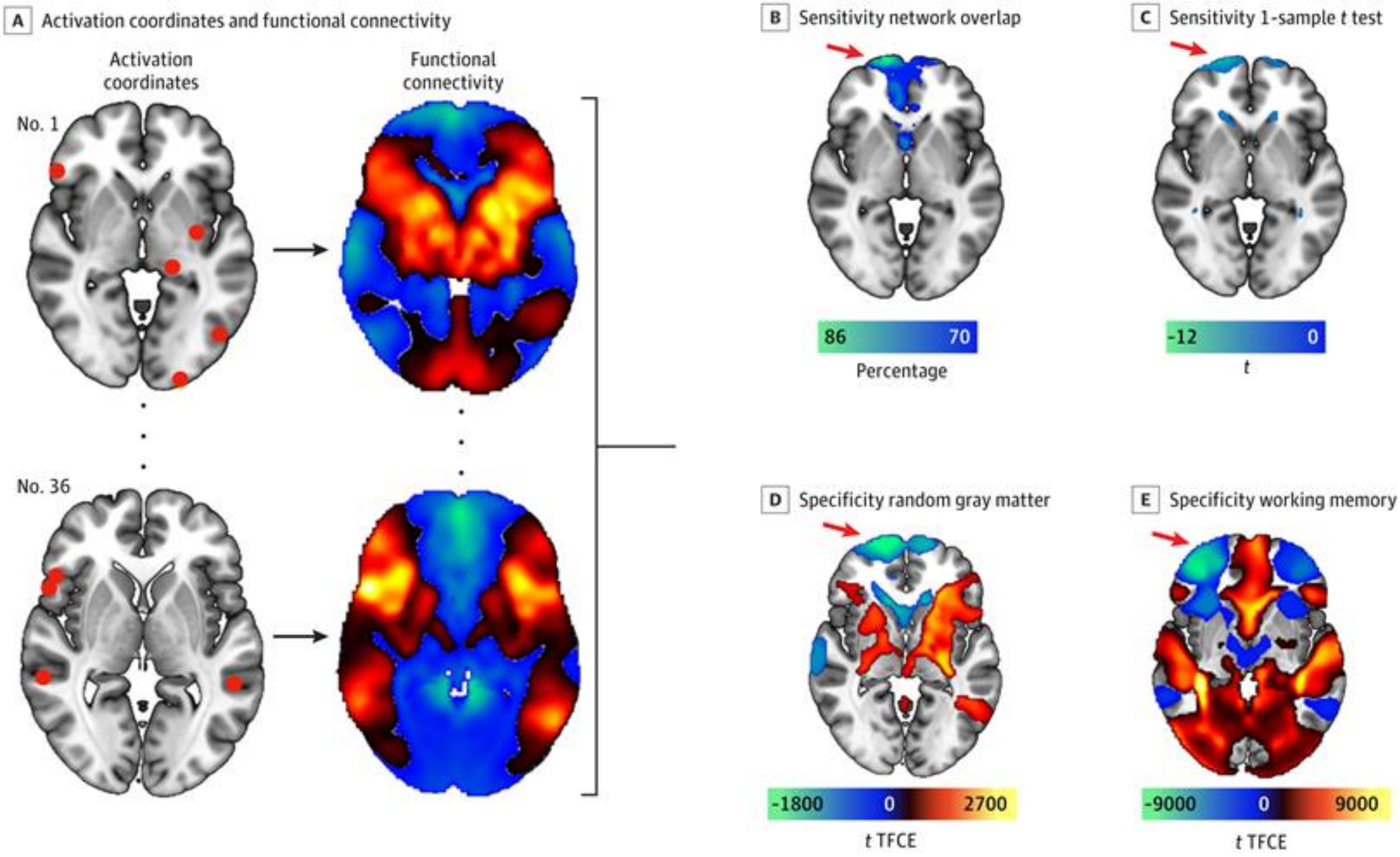
In total, 415 brain coordinates activated by creativity tasks (vs control tasks) were identified, taken from 36 individual studies involving 857 participants (pooled mean [SD] age, 24.1 [6.91] years; 461 [54%] female and 396 [46%] male). Demographic information on these participants by study can be found in eTable 1 in Supplement 1. These coordinates were highly heterogenous across different studies (Figure 1). However, 86% of these studies reported coordinates that were part of a common brain circuit (Figure 2). This circuit was defined by negative functional connectivity to the right frontal pole (Montreal Neurological Institute coordinates: $x = 20$, $y = 68$, and $z = –6$). This circuit finding was robust to methodological variation, including varying the sphere size at each coordinate (eFigure 1 in Supplement 1), running the analysis on individual-level or group-level coordinates rather than study-level coordinates (eFigure 2 in Supplement 1), varying the statistical cutoffs (eFigure 1, eTable 2 in Supplement 1), or analyzing divergent thinking and motoric creativity separately (eFigure 7 in Supplement 1). This result was also specific to creativity coordinates compared with random coordinates or to working memory coordinates (false discovery rate, $P < .05$) (Figure 2D-E). Results were again significant independent of whether we studied individual-level, study-level, or group-level coordinates (eFigure 2 in Supplement 1).

Figure 1. Neuroimaging Studies on Creativity Displaying High Heterogeneity of Results



A, Brain maps showing peak coordinates of task activation (red) from individual studies of creativity (showing results from 3 of 36 studies). B, Brain maps showing all included coordinates.

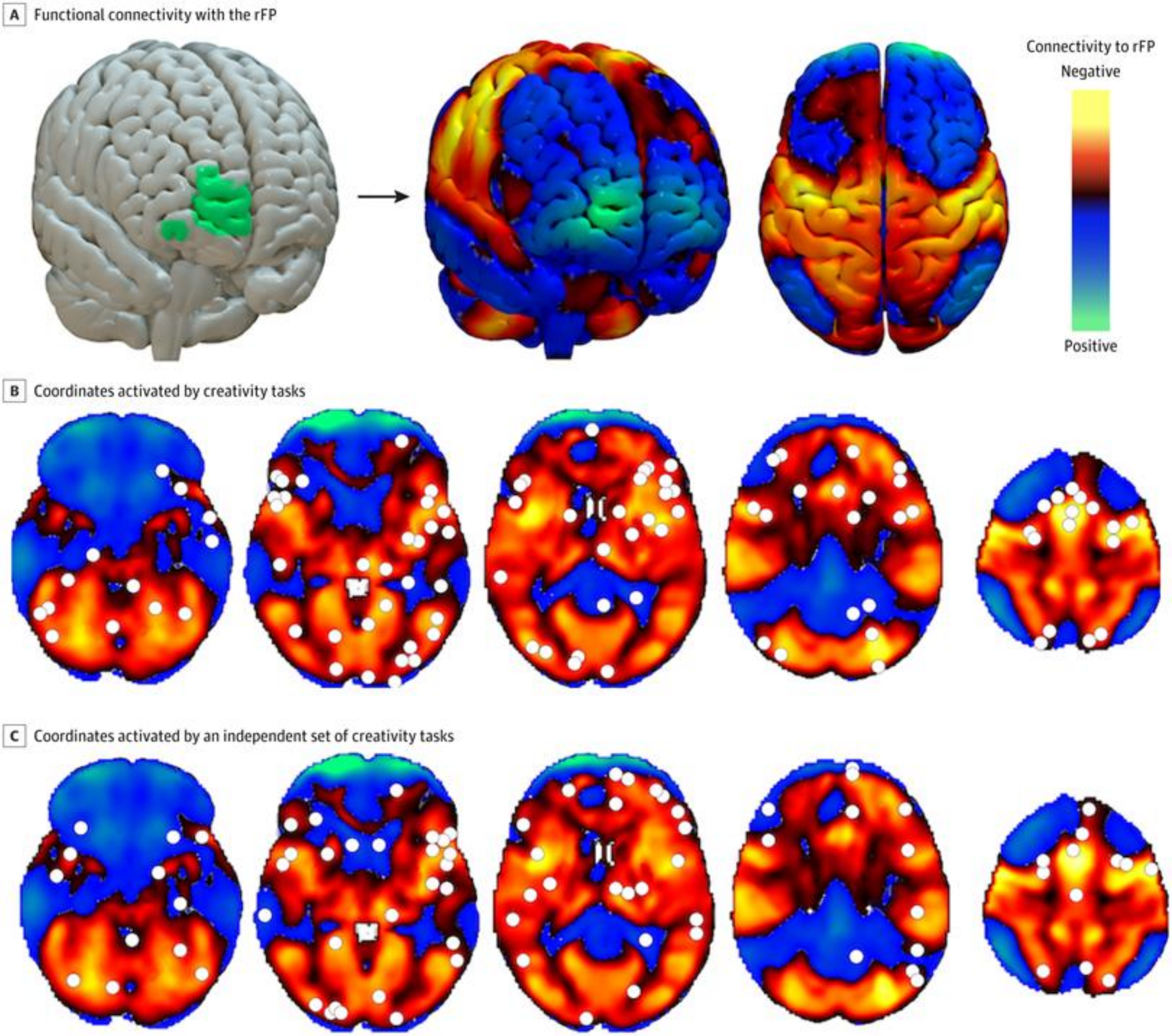
Figure 2. Methodological Overview



A, Brain maps showing peak coordinates of task activation (red) from individual studies of creativity were used as seed regions. For each study/seed, functional connectivity maps were obtained using a 1000-participants resting-state functional connectome. B, Individual connectivity maps were thresholded ($t \geq 5$) and binarized to obtain a network overlap map in which peak overlaps indicate the regions most consistently connected to the activation seeds. C, Significant results from a 1-sample *t* test at a family-wise error rate–corrected *P* value threshold of <.05. D and E, Specificity was assessed via statistical comparison with permutation analysis of linear models at 10 000 permutations against random gray matter and working memory coordinates. Reported clusters are significant at false discovery rate–corrected *P* < .05 with threshold-free cluster enhancement (TFCE). The right frontal pole (red arrows) was a consistent finding across all sensitivity and specificity analyses.

Based on these results, functional connectivity with the right frontal pole (Figure 3A) defined a brain circuit that would best encompass coordinates activated by creativity tasks (Figure 3B), but not random coordinates or coordinates activated by working memory tasks. We herein refer to this map of connectivity with the rFP as the creativity circuit. Activation coordinates from 30 independent studies (eTable 3 in Supplement 1) of creative tasks also overlapped with this circuit (Figure 3C), an overlap that was significantly stronger than expected by chance ($t_{548} = -2.75$; $P = .007$). Repeating our data-driven network mapping analysis using this independent dataset also showed peak overlap of negative connectivity in the right frontal pole (eFigure 3 in Supplement 1).

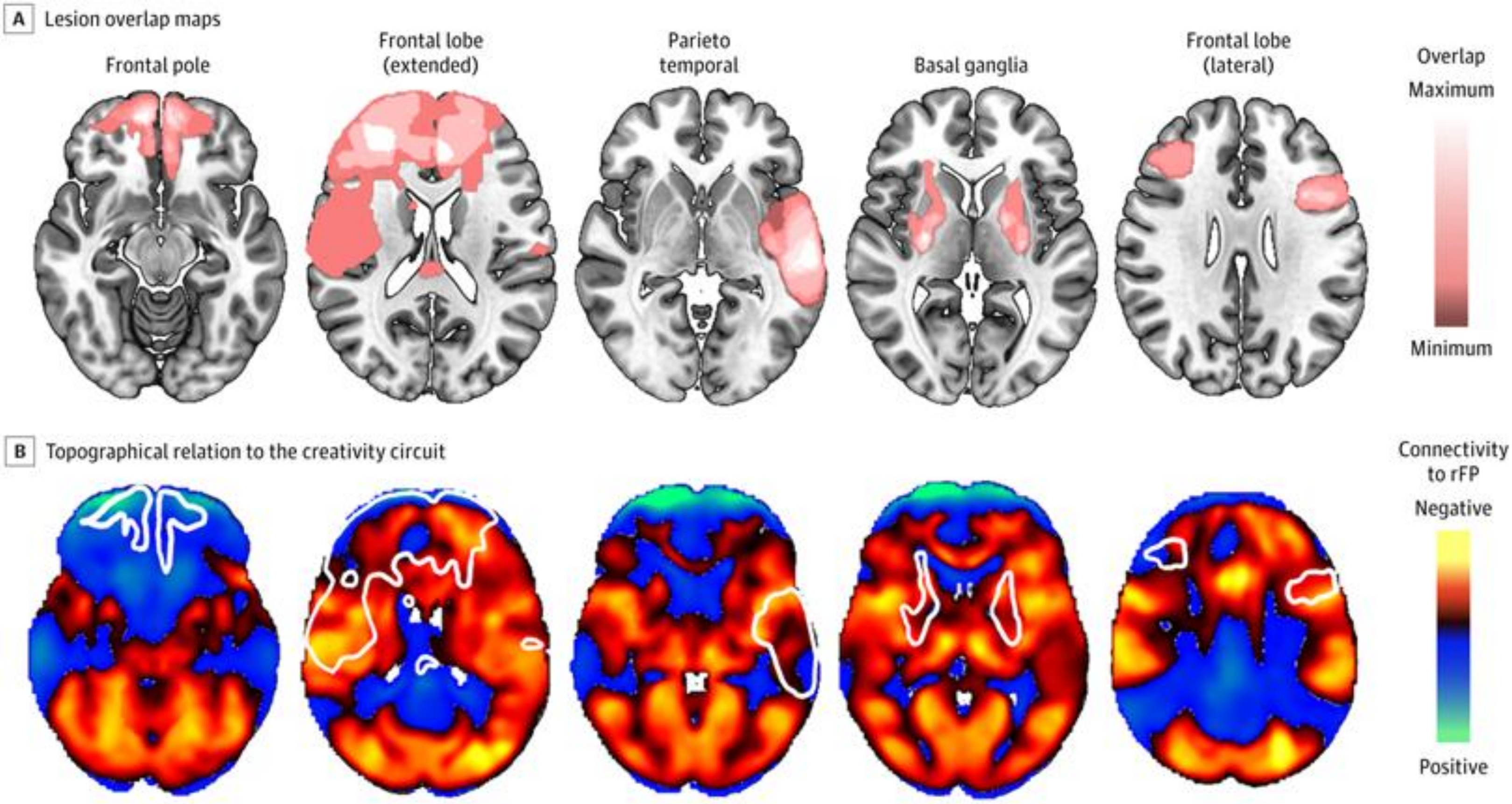
Figure 3. Connectivity of Creativity Activation Foci to the Right Frontal Pole (rFP)



A, Functional connectivity with the rFP (green, left) defines a distributed brain circuit (right) that will encompass coordinates activated by creativity tasks. B, Brain maps showing coordinates activated in 36 studies of creativity (white dots) fall within the circuit (warm colors) defined by negative functional connectivity to the rFP. C, Brain maps showing coordinates activated in 30 independent studies of creativity (white dots) fall within the same circuit.

We next tested whether damage to the identified creativity circuit from brain disease aligned with reported effects on creative abilities. First, we analyzed data from 5 groups of patients with brain lesions who previously underwent a battery of creativity tasks[31] (Figure 4A). These 5 lesion groups differed in their creative abilities, with the group having a lateral frontal lobe lesion showing the largest impairment in creativity and the group with a frontal pole lesion showing the least impairment in creativity and even an increase in creativity, although that result was not statistically significant.[31] Consistent with these findings, we found that these lesion groups intersected our creativity circuit to varying degrees ($F_{4,175} = 10.25$; $P < .001$) (Figure 4B). The group with a lateral frontal lesion intersected regions negatively connected to the rFP, similar to regions activated by creativity tasks and consistent with their lesion-induced deficits in creativity. By contrast, the group with a frontopolar lesion intersected regions positively connected to the rFP and consistent with their lack of impairment and potential increase in creativity. Post hoc *t* tests confirmed a significant difference in the degree to which these 2 lesion groups intersected our creativity network ($t_{70} = –0.32$; $P < .001$) (eFigure 4B in Supplement 1).

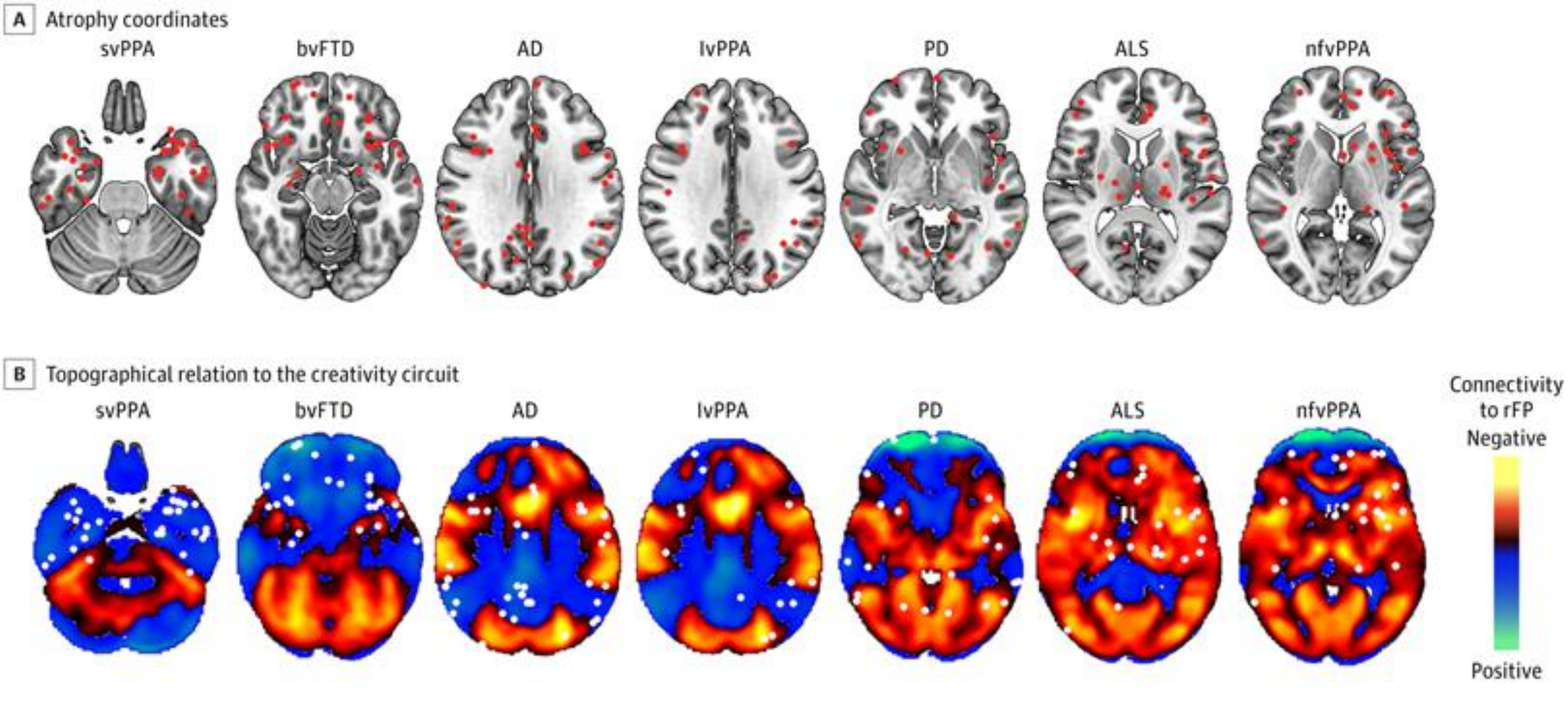
Figure 4. Lesions Affecting Creativity Task Performance and the Creativity Circuit



A, Lesion overlap maps from Abraham et al[31] showing groups of patients with lesions in different neuroanatomical locations. Patients with lesions to the lateral frontal lobe (right) had impaired performance on creativity tasks, while patients with lesions to the frontal pole (left) had higher performance. B, Topographical relation of overlap maps in (A, white outlines) to the creativity circuit. Intersection of the lesion locations with our creativity circuit was significantly associated with lesion-induced effects on creativity task performance (eFigure 4 in Supplement 1). rFP indicates right frontal pole.

Next, we analyzed the location of brain atrophy across 7 different neurodegenerative disorders, including 2 disorders previously associated with increased creativity (svPPA and behavioral variant of frontotemporal dementia [bvFTD]) (Figure 5). Coordinates of brain atrophy from these disorders differed in the degree to which they aligned with the creativity circuit ($F_{6,1295} = 6.725$; $P < .001$) (eFigure 5B in Supplement 1). Post hoc 2-sample *t* tests showed that this difference was associated with stronger alignment with svPPA than with other conditions, including nonfluent/agrammatic variant primary progressive aphasia (nfvPPA; $t_{30} = 0.18$; $P < .001$), Parkinson disease ($t_{70} = 0.08$; $P = .04$), and amyotrophic lateral sclerosis ($t_{48} = 0.11$; $P = .003$). There was also stronger alignment with bvFTD compared with nfvPPA ($t_{34} = 0.15$; $P < .001$) (eFigure 5B in Supplement 1).

Figure 5. Neurodegenerative Atrophy Patterns and the Creativity Circuit



A, Brain maps showing atrophy coordinates (red) in different neuroanatomical locations in studies on semantic (svPPA), logopenic (lvPPA), and nonfluent (nfvPPA) variants of primary progressive aphasia, behavioral variant of frontotemporal dementia (bvFTD), typical Alzheimer disease (AD), Parkinson disease (PD), and amyotrophic lateral sclerosis (ALS). Patients with svPPA and bvFTD (left) have been reported to experience increases in creativity, while patients with nfvPPA (right) have been reported to experience decreased creativity. B, Topographical relation of atrophy coordinates in panel A to the creativity circuit (eFigure 5B in Supplement 1). For svPPA, 75% of the atrophy coordinates hit regions positively connected to the right frontal pole (rFP) (cool colors). By contrast, for nfvPPA, 68% of the atrophy coordinates hit regions negatively connected to the rFP (warm colors).

Finally, we examined the intersection between our creativity circuit and the location of brain atrophy in patients with frontotemporal dementia (FTD) and new-onset creativity.[12] We found that the location of brain atrophy in these patients was negatively functionally connected to the creativity-related activation foci ($t_{35} = –2.14$; Cohen $d = –0.36$; $P = .04$). By contrast, the mean connectivity between noncreative FTD atrophy patterns and creativity activation foci was not significantly different from 0 ($t_{35} = 0.17$; Cohen $d = 0.03$; $P = .86$) (eFigure 6 in Supplement 1).

Discussion

Using network mapping of meta-analytic data, this study found that coordinates activated by creativity tasks map to a common brain circuit, defined by negative connectivity to the right frontal pole. The topography of this circuit aligned with lesion-induced effects on creativity and locations of brain atrophy associated with creativity changes in neurodegenerative diseases.

Previous neuroimaging studies of healthy participants suggest that the brain regions activated by creative tasks differ depending on the specific task or creative domain.[21] Here, we show that these heterogenous locations of brain activation map to a common brain circuit. Our findings do not preclude interesting differences across tasks or creative domains but rather suggest that despite regional differences there is a common neural substrate for creativity when assessed at the brain circuit level.

This brain circuit for creativity was defined by negative functional connectivity between coordinates activated by creativity tasks and the right frontal pole. Although the interpretation of negative functional connections remains debated,[53] negative connections are often observed between brain regions activated by tasks and other brain regions deactivated by tasks. Thus, it is possible that creativity tasks activate different brain regions but deactivate a common region in the right frontal pole. Consistent with this hypothesis, past neuroimaging research has found lower activity in prefrontal regions when individuals were asked to improvise or generate ideas compared with evaluation of ideas.[28,54,55] Other studies have implicated medial prefrontal regions near the frontopolar cortex in creative thinking,[9,56] including a study showing negative signal change when planning a creative act.[56] Because most task-based fMRI studies only report coordinates of activation, not coordinates of deactivation, it is possible that deactivation of the frontopolar cortex during creativity tasks is more common than currently appreciated.[20]

Another possibility is that the right frontal pole is a hub that defines the circuit of brain regions likely to be activated by creativity tasks, but it is not itself deactivated by creativity tasks. Either way, neuroimaging studies of task modulation cannot determine whether the right frontal pole plays a causal role in creativity. For causal inference, one must turn to other complementary data sources, such as patients with brain lesions.

Our coordinate-based brain circuit set a testable hypothesis for lesion studies: If creativity involves deactivating the right frontal pole, lesions to this region (compared with other regions) should increase creativity. Although there have been very few lesion-based studies of creativity,[57] our analysis of available lesion data supports this hypothesis. Participants with frontopolar lesions performed better than patients with lesions in other brain locations and better than healthy controls in 1 of the creativity tasks that was measured.[31]

The topography of our circuit also aligned with the changes of creativity reported among some patients with neurodegenerative diseases.[12] Specifically, our circuit aligned with increased creativity reported in svPPA and bvFTD and also aligns with a single case report of decreases in creativity in nfvPPA.[58] In their recent study, Friedberg et al[12] identified a dorsomedial occipital region to be anticorrelated with atrophy in patients with new creativity onset. While this region was not specific to creativity in our analysis, it aligned with the coordinate-based creativity circuit we found. Our result expands this finding to the effect that the visual cortex is part of the network of brain regions activated by creative tasks, but not the most sensitive and specific region involved in creative behavior across domains. Furthermore, in line with the findings of Friedberg et al,[12] atrophy patterns in svPPA most strongly aligned with the creativity circuit.

Although consistent across task activation, lesions, and neurodegenerative disease, it remains unclear why the hub of our circuit is in the right frontal pole or what role this region may play in mediating creativity. Although speculative, our results suggest that the right frontal pole may work to actively suppress creativity. Deactivation of this region during creativity tasks, damage to this region by brain lesions, or damage to this region in neurodegenerative disease may lead to a release of creativity, a process termed *paradoxical functional facilitation*.[59] Creativity is thought to involve different steps, including initial free association or idea generation followed by idea selection and refinement, the latter of which may involve analytic or self-censoring assessments.[60,61] Our circuit results may identify a neuroanatomical substrate for this latter, evaluative step. Specifically, deactivation of the frontal pole may result in decreased self-monitoring and a release of disinhibited creative output[62] or spontaneous improvisation.[54] Alternatively, the right frontal pole has been implicated in high-level cognitive control and constraining novelty seeking.[63] Transcranial magnetic stimulation inhibition of the right frontopolar cortex can result in increased novelty seeking.[64] Effects of transcranial magnetic stimulation to the right frontal pole on creativity have yet to be conducted, but this is a testable hypothesis for future research.

Our study is not the first attempt at mapping creativity to a brain circuit, and our results complement previous circuit approaches to mapping creativity.[25-29] This prior work identified correlations between brain network connectivity and creativity and in 1 case[29] validated their findings with brain stimulation. By contrast, our study identified a brain circuit that encompasses coordinates activated by creativity tasks and validated the findings using lesion and atrophy locations. Future work is needed to reconcile these different methods and results and to relate these neuroanatomical findings to an extensive literature on theoretical models of creativity.[10,61,65]

Limitations

There are several important limitations to our work. First, our maps are based on fMRI activation data that draw on a binarized creativity measure in which tasks were considered either creative or noncreative. This may oversimplify the complex structure of creativity unfolding over time or nuances of individual studies. Second, the data used to derive the creativity circuit are limited to existing literature on fMRI creativity tasks. Not all creative behavior can be studied in an fMRI scanner. Hence, there may be cases in which creativity is not negatively linked to the right frontal pole. Third, our study focused on identifying a common substrate across creativity domains, but this does not preclude important differences specific to different forms of creativity.[22,23,35] There are many different creativity domains, creativity is likely different in different individuals, and inhibiting the frontal pole is unlikely to account for the unique and diverse forms of creativity. Fourth, our approach relies on retrospective neuroimaging data. Prospective validation of the creativity circuit, especially using causal sources of evidence, such as brain lesions or brain stimulation, is needed.[66] Fifth, using working memory as a control condition allows us to increase specificity to core components of creativity, such as originality or novelty, but may mask other important contributors to creativity, such as working memory and intelligence.[24,67,68] Finally, since our method analyzes data from task-based fMRI activation studies that use different imaging processing methods, statistical methods, and study sample characteristics, potential noise associated with this heterogeneity may impact our results. However, this heterogeneity should bias against finding a convergent brain circuit for creativity.

Conclusions

The results of this study using network mapping of meta-analytic data suggest that heterogenous coordinates activated by creativity tasks map onto a common brain circuit. This circuit aligns with the effects of brain lesions and neurodegenerative diseases on creativity and provides testable hypotheses for future research.