

UNITED WISDOM

Theories of consciousness: what would allow an experimental comparison?

World and Consciousness

Andrey Zarinsh & Maria Evilkut

W-07-N03 · EN

@andrey.zarinsh @ilkutki

Theories of consciousness: what would allow an experimental comparison?

Epigraph

The unassisted hand and the understanding left to itself possess but little power.

Francis Bacon, *Novum Organum*, I, II, English text in the Joseph Devey-edited edition, P. F. Collier & Son, 1902, Project Gutenberg. The English translator was not separately established from the displayed title page.

What this article is about

When a headline announces that consciousness has been located in the brain, a short answer is tempting: which theory won? Yet theories can explain different aspects of experience, and the same word consciousness can hide different experimental questions. A meaningful comparison needs more than instruments and a large sample. It needs intelligible predictions, conditions for testing them, and a prior distinction between support, contradiction, and an unresolved result. This article considers that work through a limited comparison of global neuronal workspace theory and integrated information theory. We examine how a philosophical question becomes an experimental one, why a person's report and an activity measurement need separate explanations, and what the publicly described COGITATE comparison can teach readers. The aim is to read consciousness research without rushing to a final verdict. This is neither an exhaustive review of theories nor a clinical guide or a promise to settle the nature of subjective experience.

@andrey.zarinsh @ilkutki

I. Comparing explanations means putting them at risk

Imagine two researchers looking at one recording of brain activity. The first sees support for information becoming available across systems; the second sees an organization capable of supporting a unified experience. Each can tell a persuasive story after the result appears. But if both stories fit every possible outcome equally well, this experiment does not distinguish them. That does not establish that the theories are useless in their entirety. They may address different questions or lack sufficiently specific measurable consequences. Comparison begins when the parties identify an outcome expected under one version and less compatible with the other. Their disagreement needs a form in which an observation can change its assessment, rather than merely alter the explanatory vocabulary.

Prediction here is not prophecy about an individual. It is a claim about an observable relation under specified conditions: where, when, in what form, and with which controls a signal should appear. The scope of the conclusion matters too. A theory may concern access to content, a system’s properties, or the qualitative structure of experience. Testing one consequence does not test every part of its explanation. A major theory’s name in a headline often raises expectations beyond what the actual task can establish. Readers can begin with a simple question: what, specifically, could have turned out differently here? The answer reveals whether predictions genuinely collide or merely redescribe an already familiar observation.

There is a human side to this work. A theorist may have invested years and see a new result as the continuation of a strong argument. An opponent may await refutation with equal commitment. An inspectable agreement need not pretend that interests disappear; it gives them less power to change the rules after the fact. A hypothetical agreement can specify analysis, data boundaries, and treatment of uncertainty. Agreement itself does not guarantee truth: it can be poorly designed. Its merit is that substantive disagreement becomes easier to assess. Another researcher can see which grounds were promised, what was actually measured, and which revision follows from a particular mismatch.

II. First establish which consciousness question is being asked

One word can gather too many tasks. ‘The person is conscious’ concerns a state; ‘they are aware of the red object’ concerns a content; ‘they can report it’ concerns a response ability. These formulations are related but not interchangeable. The 2011 review by Dehaene and Changeux explicitly addresses conscious access: how information becomes available as a reportable subjective experience. Read that boundary together with the model. An experiment concerning the content of visual perception does not automatically answer every question about sleep, pain, or deep meditation. Specifying the domain does not diminish the investigation. It establishes which particular step was made and what a subsequent study would still need to examine.

Consider a hypothetical conversation about music. A person hears a melody, notices an instrument change, and explains why they like it. Three nearby questions arise: did they experience the sound, detect the change, and construct an explanation? An answer to the last does not exhaust the first two. Research would need to establish which difference counts as the relevant observation, how it is reported, and which aspect of behaviour is compared. Asking simply whether consciousness occurred does not make the word a single measuring instrument. Difficulty defining the task does not mean there is no subject matter: the person actually reports experience. It means deciding how this report participates in testing this explanation.

Comparing theories therefore requires a shared target without forced agreement on every definition. Parties can agree to investigate a particular perceived image while disagreeing about why it is experienced. Name the agreement and its boundary. If one side treats report as an indicator of access and another seeks properties of experience itself, that difference should enter the reasoning before results appear. Otherwise the expected victory becomes a dispute about what victory meant. Scientific work starts here with intelligible links between terms and procedures. A good definition does not replace measurement, but it helps prevent demands that a measurement answer a question never built into the task.

III. Two theories are not merely two metaphors

Global neuronal workspace theory, GNWT, links conscious access to information becoming broadly available across systems. In the Dehaene and Changeux account, this is a proposal about neural organization, not a little observer inside the head watching a screen. Workspace is a convenient term rather than a physical room. The limited distinction needed here is that locally processed information and information available to several processes are not treated as the same state within this model. Detailed neuroanatomy and mathematical implementation are outside this article. Broad availability is presented as a theoretical proposal, rather than a proven explanation of every qualitative feature of experience or every possible conscious system.

Integrated information theory, IIT 4.0, begins with properties of experience formulated by its authors and proposes a physical account in terms of a system's cause–effect organization. Integration here does not simply mean many messages or a well-coordinated team. It is a technical concept within a particular theory. The authors connect experience with an irreducible cause–effect structure and develop a formal account of it. These are IIT's claims, not truths established by unanimous research agreement. The 2023 paper has no new human sample: it presents a theory, formal framework, and examples. Treating it as an additional human experiment would artificially multiply empirical support.

Comparison becomes more interesting when both theories stop being translated into the convenient word connectivity. A network may distribute information widely while the question of its own causal structure remains separate. A theory may specify a structure without a particular instrument measuring everything its formalism requires. Between a broad proposition and an observable signal lies a bridge: which features of a real brain are acceptable indicators of the proposed mechanism? That bridge needs an argument from model proponents and examination within the study. Without it, an attractive term acquires too many jobs. We may think explanations were compared while actually comparing two equally appealing narratives about connections.

IV. Predictions must diverge observably

A helpful experiment needs more than an expectation of brain activity. If both accounts permit activity, its presence creates no decisive difference. A more specific consequence is needed: the distribution of information, persistence over time, or a particular relation between measured sites. A detailed prediction binds theory to conditions. It may concern stimulus onset, continued presentation, or disappearance. Crucially, it belongs to the version being tested rather than to a reader's reconstruction from a general impression. Theories develop; a description from one version cannot quietly replace an agreement made for another. The history of a change also belongs in an intelligible account of the result, rather than vanishing behind an unchanged theory name.

Take a hypothetical example, not a description of particular COGITATE data. Model A expects a brief signal when an image appears, whereas model B expects a specified signal to continue throughout viewing. A researcher varies duration and measures the temporal profile. If the signal lasts longer with a longer presentation, there is information for comparison. Yet another step is needed: can the instrument distinguish brief from prolonged processes, does processing smooth the data, and could the result reflect preparation of a response? A prediction gains strength through intelligible control of alternatives, not through a confident A or B placed beside an attractive graph. The graph represents measured and processed observations, not the explanation itself.

Ambiguous outcomes deserve a prior description too. Measurement may lack sensitivity. Results in one region may fit a prediction while results elsewhere conflict. Previously acceptable methods may yield different findings. 'Unresolved' then conveys definite information about the test's limits. It does not mean no work was performed, and it does not authorize automatic selection of the preferred theory. Distinguish the primary plan from subsequent exploration as well: an interesting pattern discovered after analysis can motivate a new prediction, but cannot retrospectively become predicted. That distinction protects the value of both discovery and testing, instead of converting every unexpected finding into a success supposedly anticipated from the start.

@andrey.zarinsh @ilkutki

V. The participant’s report belongs to the experiment

Subjective experience does not arrive as an image on a shared screen. A person can report seeing or hearing, and a researcher compares that report with behaviour and measurement. Reporting itself requires operations: understanding instructions, retaining content, selecting a response, and moving or speaking. This is a logical distinction, not suspicion of participants’ honesty. A signal near the report raises a question: does it concern experience, preparation to answer, or several related processes? ‘The person said it, so consciousness has been located’ cannot settle that. Nor can a report simply be discarded because it is subjective; doing so may remove the needed connection between the measurement and the research question.

In a hypothetical task, one object requires no response and another requires a button press. Comparing these conditions may reveal a signal difference related to the task, rather than solely to experienced content. Researchers must specify what changes and what remains constant. A control condition helps, but the word control does not guarantee a good comparison. What matters is similarity on relevant properties and the side differences that remain. The participant may expect the target differently, watch for it, or prepare a movement. These possibilities do not refute a study in advance. They become questions for a design that should permit a more definite interpretation of the observation and its relation to conscious experience.

The COGITATE first-experiment documentation explicitly includes task relevance and presentation duration. It specifies three relevance levels and three durations: 500, 1,000, and 1,500 milliseconds. This illustrates the structure of a question; it is not our replication of the experiment. We do not claim that one design removes every report-related difficulty or suits every form of consciousness. A useful reader’s step is to ask how far measurement separates experienced content from the requirement to act upon it. If a report of results does not show that connection, the need for it has not disappeared. It remains a boundary of the explanation available to the reader.

VI. What collaborative disagreement contributes

The 2025 COGITATE Consortium paper compared IIT and GNWT predictions. In the abstract actually read, 256 participants viewed suprathreshold stimuli of varying duration, with functional MRI, magnetoencephalography, and intracranial electroencephalography used. Design and interpretation were preregistered with theory representatives. Reported findings partly fitted predictions and substantially challenged some tenets of both theories. This is one study, not proof that comparison is now always straightforward. Our reading covers the abstract and selected figure captions, alongside separate project documentation; it does not claim full reading of every method or supplement. This stated scope limits the level of interpretation offered here.

The point of this collaboration is not simply reconciliation between scientific schools. Parties may continue to disagree while having to specify which outcome would trouble their version. When a common task is clear enough, surprise becomes shared: positions encounter an agreed set of conditions rather than each team separately collecting convenient signs. Agreement still leaves questions about measurement adequacy, construction of indicators, and later interpretation. Collaborative work is valuable as a way of making those questions visible. Collegiality is not independent evidence of truth, and comparing these accounts does not establish that every important alternative theory was included. The comparison must retain the boundaries of its participants and chosen task.

For a nonspecialist, distinguish three voices: the prior expectation, the observed result, and the theorist’s present explanation. They may differ. A later explanation should not be read as though it had always been the original prediction; neither should all revision after new data be forbidden. Revision is legitimate when its status is clear. It may retain part of a mechanism, abandon a particular consequence, or propose another measurement relation. The next step is then examination of the revised prediction, not declaring the earlier version unconditionally successful. Work accumulates through traceable change, allowing readers to understand what the updated theory explains better or differently and which proposed improvement remains to be tested.

VII. Not detected does not automatically mean absent

A negative result needs as much precision as a positive one. When an expected signal is not found, ask which signal was sought and whether the method could reliably notice it. No detection in an analysis differs from absence of the phenomenon. Yet that distinction cannot become permanent rescue for every theory, or no outcome could constrain it. A specific measurement problem tied to the conditions is needed, rather than the general statement that science is imperfect. If conditions were previously accepted as adequate for testing a prediction, a negative outcome creates a substantive difficulty. A new objection remains legitimate when supported and openly distinguished from the earlier agreement.

The strength of the claim matters too. A model saying a process is always necessary takes a different risk from one expecting it only in a particular task. If a limited finding conflicts with the former prediction, ‘always’ cannot quietly become ‘sometimes’ while retaining the previous status. A narrowing can honestly be proposed and then examined for whether it still explains what the theory is meant to explain. A narrower version may become more precise at the cost of scope. That cost is not a disgrace; it belongs to research. Readers benefit from seeing what was lost and gained rather than a story of endless victory in which every wording change confirms an allegedly unchanged idea.

Comparison can also produce a mixed outcome: one expected feature appears and another does not. Mixed findings should become neither complete failure nor complete confirmation. A good report retains a map of particular predictions. Readers may lack the expertise to assess every analysis, but they can notice when a headline extends a partial finding into a universal conclusion. ‘Which theory won?’ then gives way to a more useful question: ‘Which commitments survived this test, which are challenged, and what remains open?’ This does not evade evaluation. It makes evaluation more precise and connects it to observations rather than requiring a final scientific allegiance on the basis of one collection of results.

VIII. Instrument, indicator, and explanation

An instrument measures a particular physical relation; it does not deliver the entire human question in finished form. Measurement becomes an indicator, the indicator is connected with a proposed process, and the process with the theory’s explanatory target. Each transition can be reasonable, but it should be named. A program distinguishing stimulus categories from a signal has not thereby read the whole subjective experience. A signal related to presentation duration still needs an account of that relation. These boundaries do not deny neuroscience. They help identify what a particular method does and which additional comparison is needed for a stronger conclusion, making the actual result more informative rather than less valuable.

Information is especially prone to crossing boundaries unnoticed. In ordinary speech it concerns a message's meaning; in a measurement it may concern distinguishable states; within a theory it has a special definition. These uses can be related, but their relation needs demonstration. An algorithm distinguishing image A from image B has not automatically measured integrated information in IIT's particular sense. General signal exchange does not automatically establish a workspace as specified in the tested GNWT version. Readers need not solve the mathematics to notice this. Look for where investigators explain how the indicator corresponds to the concept, and acknowledge honestly when that explanation is not present in the material actually read.

In Novum Organum I, II, Bacon compares assistance for understanding with instruments for the hand. His argument belongs to the history of inquiry, not contemporary instrumental neuroscience. A bounded comparison is nevertheless useful: concentrated thought alone is insufficient; means that make a question examinable matter. Today those means include definitions, analysis procedures, and retained grounds of inference as well as apparatus. This is our methodological interpretation, not a claim that Bacon anticipated MRI or protocol registration. A historical text helps formulate the task. A contemporary experiment must address it through its actual methods and observations rather than borrow past authority in place of its own examination.

IX. An experiment and the nature of experience

Research can clarify a relation between observed activity and experienced content while leaving philosophical disagreement about what the relation means. That does not authorize treating every metaphysics as equally established. An empirical conclusion and a broader interpretation simply have different scope. Supporting a strong ontological claim requires another argument explaining the transition. Difficulty measuring does not prove that consciousness is independent of the brain; an observed relation does not automatically resolve every philosophical problem. Readers can remain interested in both questions without requiring one experiment to settle them together. A precise account of the accomplished step makes clear which work still lies ahead and prevents remaining questions from erasing what has actually been learned.

A cultural or religious account of experience may pose another question: how to live with it, understand its meaning, or pursue a discipline of attention. Similar words do not turn that account into a neural theory. This article does not review particular religious teachings, so it assigns them no shared position and adds no unrelated quotations to fill a list of traditions. The boundary is substantive: personal meaning and scientific explanation can coexist in one life while their grounds are assessed for their respective tasks. A particular physical mechanism

claimed on behalf of a tradition requires appropriate support of its own. Respect for a worldview does not turn every use of it into an experimental finding.

Philosophical doubt should not substitute for particular criticism either. ‘Consciousness is too complex’ does not identify an error in an analysis. ‘The brain does everything’ does not show which model explains a result better. Between these slogans lies research: refining conditions, comparing consequences, examining alternatives, and preserving unresolved questions. A public article can end with that definite structure without manufacturing a final answer. For an adult reader, incompleteness need not mean disappointment. It makes the structure of knowledge visible: what rests on observation, what is proposed as explanation, what has been tested, and what still lacks sufficiently discriminating support. The remaining difficulty can then guide a next question rather than become an all-purpose excuse.

X. Reading a claim that a theory has won

When an arresting headline announces an explanation of consciousness, reconstruct the path from name to conclusion. First identify the domain: state, content, or access. Find the tested predictions and actual conditions, including people, task, and methods. Then distinguish observed features from interpretation. This order does not require joining a school in advance or suspecting every scientific article. It makes trust in a particular result possible without unsupported extension to everything. An abstract can be an honest beginning; a detailed strong conclusion based on unread methods is not yet the reader’s own finding, even when the paper appears in a prominent journal. Access and reading scope remain part of the assessment.

Ask about revision too: which part of the explanation became less secure or gained support? If the answer is always ‘the theory was confirmed’, regardless of mixed findings, reporting has lost its discriminating function. If it is always ‘everything was disproved’, examination also disappears behind a dramatic gesture. Scientific rivalry need not end with one winner. It can reveal shared assumptions needing clarification, ambiguous indicators, or model components that survive different tests. These are possible kinds of outcome, not prescribed endings for every experiment. Reporting them needs connection to a particular study and its limits, instead of turning complexity into another attractive general slogan that cannot be challenged by any observation.

Return to the two researchers facing a recording. Their substantive encounter begins neither when they find the same metaphor nor when one falls silent. It begins with the question of which differences they are genuinely prepared to test. An instrument, clear procedure, and open analysis history help make that difference a shared object. After testing, explanations may change unequally, and part of the question may remain open. This is a modest but real step toward understanding human experience: instead of choosing an attractive story, we can

compare stories' commitments to observation. A theory becomes scientifically interesting because it explains something definite and allows investigation of where its explanation fails, rather than because it finally reassures us.

What this means in practice

Take one accessible report of consciousness research and make a five-line map: question; theories and versions; divergent predictions; observed outcome; limit of the conclusion. If you read only the abstract, record that. Do not fill in from memory a sample size, method, causal mechanism, or successful outcome absent from the text.

Choose a headline verb such as proved, refuted, discovered, or explained. Find the paper's sentence meant to support it and compare the scope. If the study examined visual content in a particular task, do not extend its outcome to every experience. If findings are mixed, preserve differences between predictions. A stronger assessment requires the methods and appropriate competence; access to a document does not mean every analysis has been evaluated.

Write your own limited account without a victory label: 'Under these conditions, the researchers found a result challenging this prediction; another question remained open.' This example of form does not replace the chosen paper's actual data. In conversation with someone favouring another theory, consider what next outcome would change your assessment. This is voluntary reading and discussion practice, not a tested means of eliminating bias or an experiment on your own consciousness. Its value to you can be assessed through whether the question, grounds, and remaining boundary became clearer.

Key quotation

> *But experience is by far the best demonstration, provided it adhere to the experiment actually made, for if that experiment be transferred to other subjects apparently similar, unless with proper and methodical caution it becomes fallacious.*

Francis Bacon, *Novum Organum*, I, LXX, first sentence in the Devey-edited 1902 English edition. A historical claim about inquiry, not an established neuroscientific proposition.

Connections between traditions

1. **GNWT: access / IIT: experiential properties.** Both propose relations between consciousness and physical organization, while their starting targets and definitions differ. Broad information availability is not identical to irreducible causal structure. Grounds: Dehaene–Changeux 2011; Albantakis et al. 2023.

2. **Bacon I, II: assistance for understanding / COGITATE: agreed procedure.** The similarity concerns means of inquiry rather than one mind's confidence. Historical inductive philosophy neither predicts modern design nor replaces observations. Grounds: Bacon and the COGITATE project account.

3. **Bacon I, LXX: experimental boundaries / IIT–GNWT comparison 2025: tested-prediction boundaries.** The comparable distinction is between a result and extension beyond its conditions. The modern investigation concerns a particular human task; Bacon offers methodology rather than a neural theory.

4. **IIT 4.0: formal framework / COGITATE: measurable consequences.** Both levels need clear indicators, but theoretical mathematical structure and a chosen empirical signal are not automatically the same. Their connection requires justification. This article does not claim direct measurement of the entire IIT formalism.

5. **GNWT: reportable access / experimental task relevance.** Both require distinguishing content from operations using it. Removing a task does not yield transparent access to all experience, and reporting is not merely fabrication. Grounds: Dehaene–Changeux review and COGITATE project documentation.

Primary sources

1. Francis Bacon. Novum Organum, I. Joseph Devey-edited English edition,1902; primary philosophical text, II/LXX quotations and XIX/XXV/XLVI context.

2. Dehaene, Changeux. Neuron, 2011,70:200–227. GNW proponent review; abstract and selected PDF pages concerning access and the model, rather than every underlying experiment.

3. Albantakis et al. IIT4.0, PLOS Computational Biology,2023. Primary formal theory paper; introduction/postulates, with no new human sample.

4. COGITATE Consortium. Nature,2025,642:133–142. Primary experiment; our reading covers abstract and selected figure captions, not full methods and supplements.

5. COGITATE Experiment1 overview. Primary project documentation of task conditions; not an independent replication.

Tags

consciousness, consciousness theories, experiment, GNWT, IIT, COGITATE, conscious access, subjective experience, content, state, report, prediction, preregistration, theory testing, measurement, indicator, method, control condition, negative results, mixed findings, Bacon, Novum Organum, Dehaene, Changeux, Albantakis, interpretation, inference scope, philosophy of inquiry, causal structure, collaborative testing

About the authors

Andrey Zarinsh is the author and founder of United Wisdom. He explores inner transformation, the unity of traditions, intercultural dialogue, and contemporary spiritual discipline. Instagram: @andrey.zarinsh.

Maria Evilkut is a co-author of United Wisdom. Alongside her research, she seeks to test these ideas in practice and observe how principles of unity and mutual understanding work in real life between people and cultures. Instagram: @ilkutki.