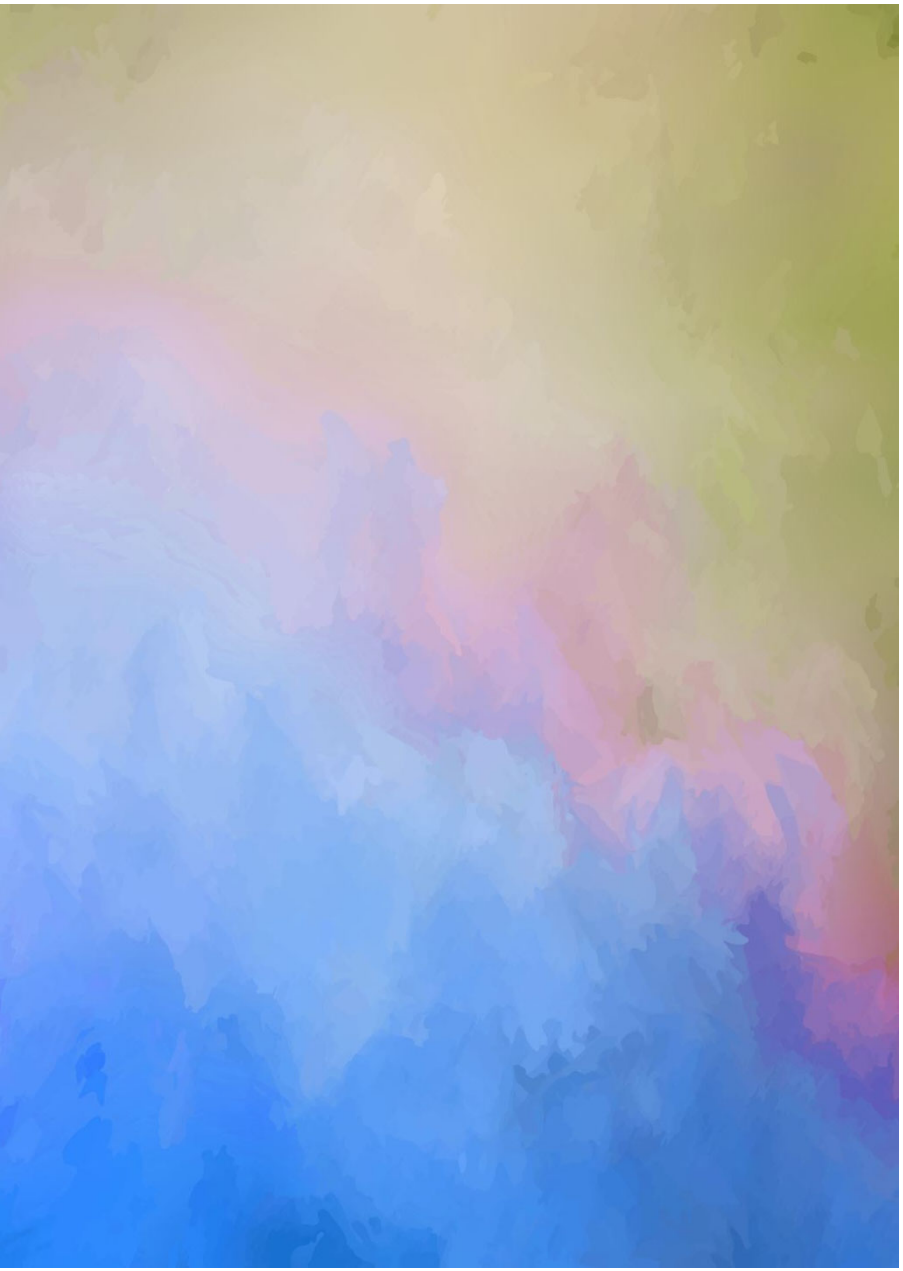


A Framework for a testable theory of consciousness

Yoshiyuki Ohmura (The University of Tokyo)

Yasuo Kuniyoshi (The University of Tokyo)

C-21(THU-2), October 15, 5:00
Dockside Room, TSC 2026

What is the difference between mental representation and information?

- Non-physical: But what is the definition of “physical”?
- Objectively indescribable: But cognitive science has demonstrated that cognitive functions can be realized by treating mental representations as information.
- Subjectivity: It is the most natural reaction. Can subjectivity be defined objectively?

Yes: If a subject can be defined objectively.

Consciousness without a subject

- Background

Much of contemporary theory of consciousness accepts the concept of consciousness without an experiencing subject.

In many theories, the subject is not defined.

Can consciousness without a subject explain the characteristic of subjectivity?

Establishing the Concept of Consciousness Without a Subject

- It appeared in a book by Parfit[1984] that aimed to eliminate the concept of the self. He adopted an interpretation of split-brain cases as involving two streams of consciousness (Puccetti, 1981).
- Dennett [1991] argued that introducing an experiencing subject leads to the homunculus fallacy.
- Crick [1994] argued that the homunculus inside the brain is legitimate explanatory target in the science of consciousness.

Link to Reference

The Historical Role of Consciousness Without a Subject

This concept encouraged a focus solely on the computational aspects of cognition and contributed to the advancement of cognitive science.

But, since this concept does not account for subjectivity, it should not be equated with mental concepts.

When Chalmers proposed the “hard problem,” this concept was never reexamined.

If consciousness emerges without a subject, it contradicts the subjectivity of consciousness.

Confusion Between Concepts (Subject, Subjectivity, Subjective Experience)

Objectively describable

- Subjectivity (if a subject can be described objectively)
- Subject (legitimate explanatory target)

These concepts can be introduced in theories of consciousness.

Objectively indescribable

- Subjective Experience
- Qualia

To make a testable theory, we cannot introduce these concepts in theories of consciousness.

A Subject-First Approach

- Subjective experiences influence behavior and verbal reports. This implies that subjective experiences have the corresponding physical correlates of consciousness (PCC).
- A testable theory of consciousness must consist solely of concepts that can be described objectively.

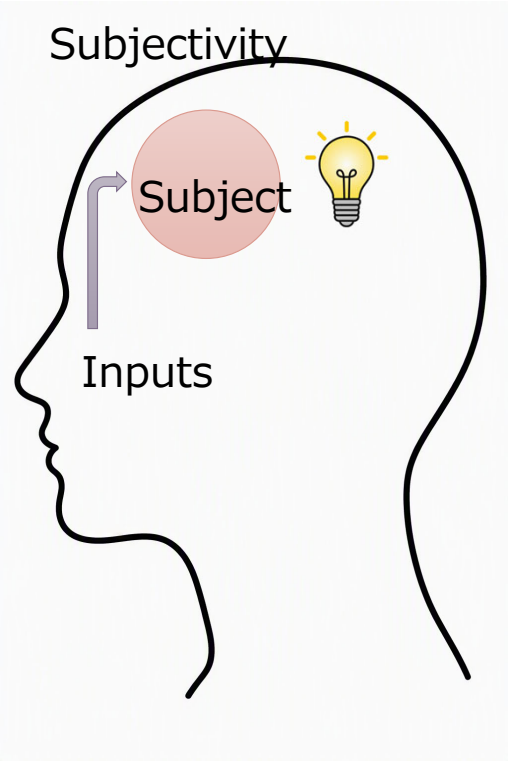
A Subject-First Approach

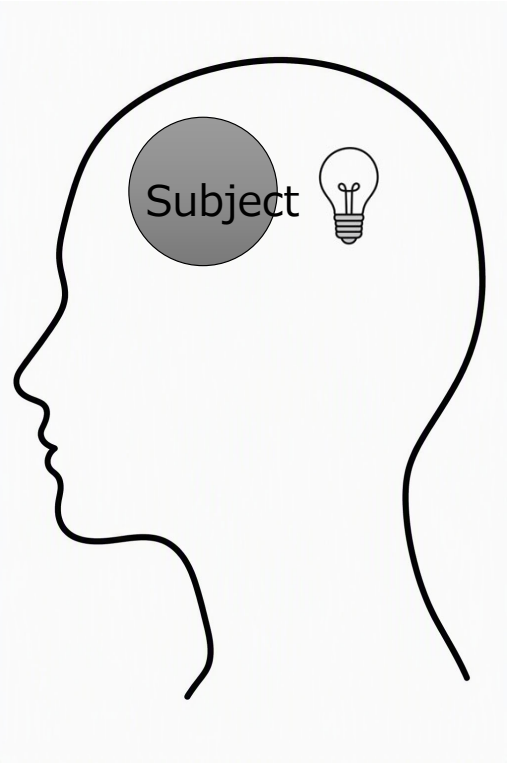
- Describes the subject of experience objectively.
- Includes the concepts of Subject, Subjectivity, and PCC.
- Does not theorize about the relationship between PCC and subjective experience.

Link to Reference

Simple and Intuitive Explanation

- Whether conscious or not is determined by functional realization of a subject.
- Inputs available to the subject is conscious perception.
- The only remaining issue is the definition of “Subject.”





Conscious

Unconscious

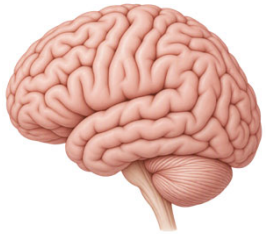
A model of mental causation

- We define mental concepts as subjective.
- Definition:

Mental Causation is a causal effect from PCC to other physical components.

Because PCC can be objectively described, mental causation also can be described objectively.

PCC inside the brain → Mental Causation



If a subject can be defined, PCC and mental causation can be defined.

Comparisons

- **Subject-First Approach**

Realization of the subject and availability of inputs are separable.

PCC, subject, subjectivity and mental causation can be defined.

Legitimate explanatory target is a subject.

- **Conscious-Centered Approach**

All conscious related phenomena must be explained by multiple aspects of consciousness [see Bayne, 2016].

It is difficult to explain the differences between consciousness and information without taking subjectivity into account.

Characterizing the experiencing Subject

- It is related to the input that constitutes the contents of consciousness.
- Whether functionally realized or not depends on the state of system.
- Consciousness is transient, but the subject is relatively persistent.
  - We focus on a causal and physical structure, rather than informational or statistical structure.
- It cannot be realized in general physical systems, but it can be realized in the brain.

We assume that “persistent asymmetry” is an important characteristic that needs to be explained.

Persistent asymmetry cannot be explained by general causal relationships.

- Directional causation cannot account for causal effect of the subject.
- Interactions lose their asymmetry.
- A special type of causal mechanism is required.



Intrinsic Causation

- Intrinsic causation is a causal relationship between two systems that share a set of physical substrates.

Definition:

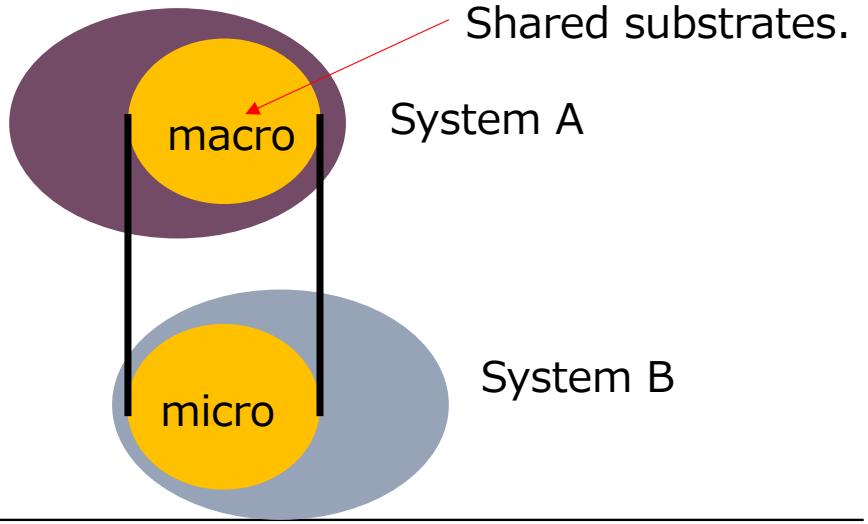
1) $C_A \rightarrow C'_A$ represent a causal transition occurring within System A.

2) E_B represents an effect occurring within System B.

3) Shared Substrate: The substrates realizing C'_A and E_B partially overlap: $S_{C'_A} \cap S_{E_B} \neq \emptyset$.

4) The occurrence of E_B depends on the occurrence of C'_A.

5) Shared substrates have different functional roles across systems.



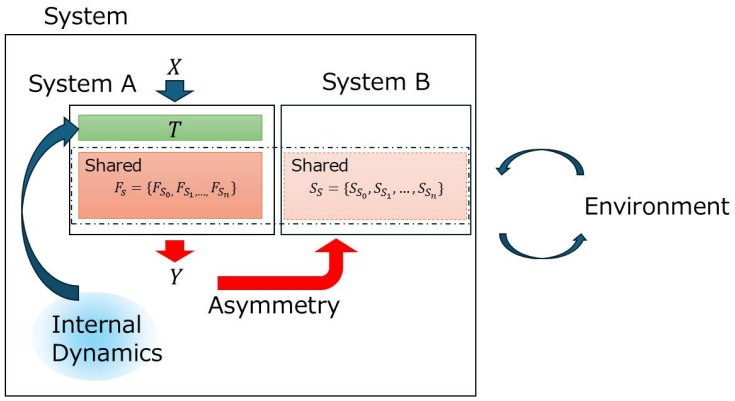
Intrinsic Causation and hierarchical causation

- Intrinsic causation is similar to hierarchical causation, but there are significant differences between the two.
- Differences:
  - Hierarchical causation generally does not assume the existence of any substance other than shared physical substrates.
  - Hierarchical causation assume that the ultimate origin of cause should be attributed to a higher level itself in the hierarchy. However, to explain persistent asymmetry, the origin of cause is not important.
  - **Such relaxations are crucial to realize intrinsic causation in physical system.**

Realization conditions for intrinsic causation

The paper linked here explains the conditions necessary for intrinsic causation.

The abstract mechanism of awareness is naturally derived from the method of realizing intrinsic causation.



Y. Ohmura and Y. Kuniyoshi
Subject-Centered Approach in Consciousness Science, Zenodo, 2026

Summary

- To construct a testable theory, a theory consists of objectively describable concepts alone.
- If a subject can be defined objectively, PCC, subjectivity and mental causation can be theoretically defined.
- Objectively indescribable concept is qualia!
- Subject-First Approach is a promising method for a theory of consciousness.

Y. Ohmura and Y. Kuniyoshi
Subject-Centered Approach in Consciousness Science, Zenodo, 2026