

The Brain & Consciousness

Ref: Baars, Preface, pp.6-7, 27-31, 40-61 (esp.55-59), 62-66, 73-78

- We'll first borrow mostly from Baars concerning brain details; Baars ch. 1-2: The close correlation between subjective experience and observable brain processes *explains subjective experience*
- Then we'll go back to Blackmore, ch. 2: overwhelming evidence that *consciousness is determined by brain physiology*

“Theater of mind” metaphor

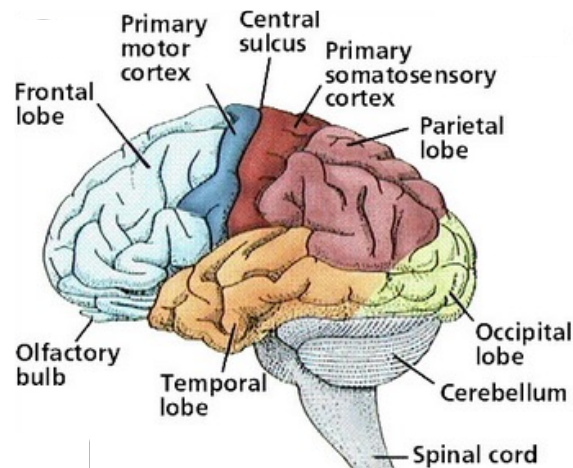


An internal “screen”, and a “self” observing it → infinite regress!

The “screen” is not so far-fetched (cf. V1-V5 in visual cortex), but the processing is specialized, not a “self”!

Baars’ “Theater of Consciousness” theory:
a theater is exactly the right metaphor

But first: brain structure and function.



The lobes of the brain

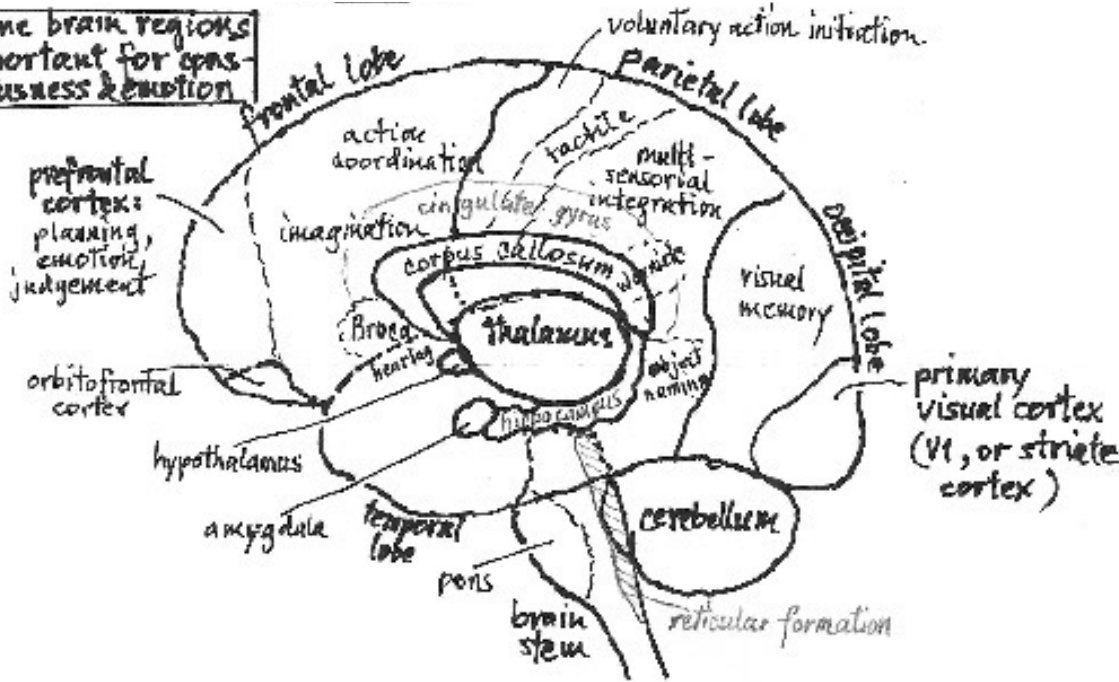


Associated functions

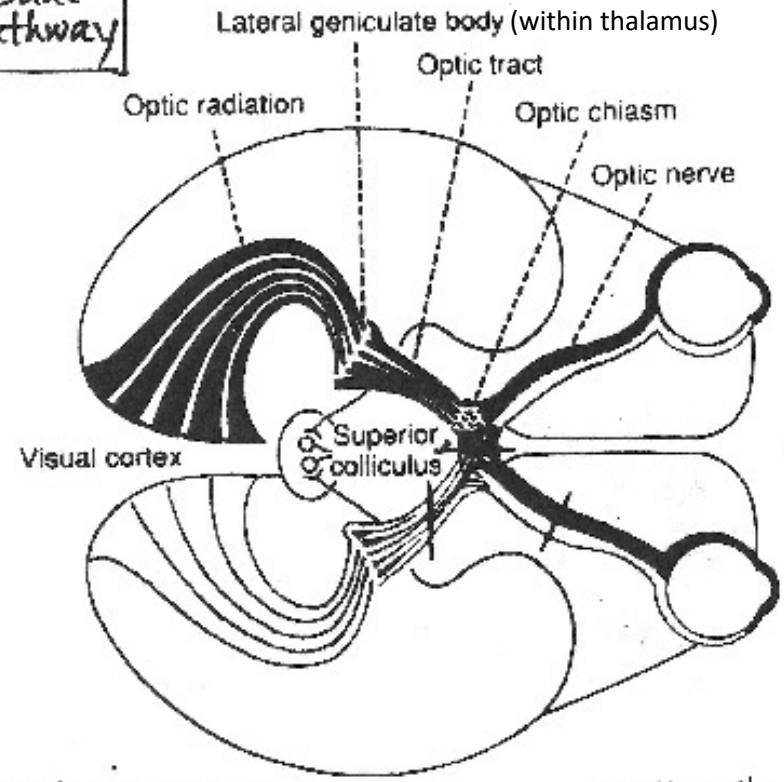
Brain Systems Involved in Consciousness

Brain interior

Some brain regions important for consciousness & emotion



Visual pathway



Main ideas to take from Baars preface & ch. 1

Global Workspace: neural system that displays conscious events, especially sensory

Explanatory Gap, acc. to Baars: Close agreement of subjective & objective phenomena closes the gap (Big Switch?!)

PET scans, inserts 1-4: diminished brain activity as task becomes “automatic”

Brain facts, overall: 86 billion neurons, 10,000 dendritic inputs per cell;
7 “steps” from one cell to any other;
Neural layers, columns, maps, clusters, routines, “cables”;

Consciousness function: the “publicity organ” of the brain, providing global Information access/dissemination, coordination, control

Systems essential to C: reticular formation; thalamus; ILNs; reticular nucleus;
40 hz cycle (Llinas & Crick);

Karen Quinlan – both thalami deprived of blood

Main ideas to take from Baars ch. 2

Theater model: the “stage” of consciousness provides limited capacity, vast access; Spotlight (attention), audience (unconscious processes), director (?)

More brain facts: 6 cortical layers

- outermost gets most direct perceptual input;
- progressively more abstract layers;
- innermost layer most closely associated with consciousness;
The layers consist of narrow adjacent columns; each seems to do similar 6-layer processing (but is also cross-connected);
- again note especially the role of the thalami as relay stations & in consciousness;
- cabling: corpus callosum, connecting the hemispheres, and heavy front-to-back cables, underneath hemispheres;
- hippocamp{us,i} – multiple memory functions

Reasons for limited consciousness: better learning through undivided attention; slow speed of neurons might impede “multi-threading/processing”?

Biofeedback & learning: can learn to control any neural system, given perceptible feedback

Brain anatomy & function (see earlier images)

See more detailed notes at

<https://www.cs.rochester.edu/~schubert/191-291/lecture-notes/brain-anatomy-and-function>

Terms to learn: Nucleus (as it applies to assemblies of neurons), gyrus/gyri, sulcus/sulci, ventral, dorsal, medial, inferior, superior, frontal, posterior, lateral, reticular, orbital, soma

Major brain lobes & functions:

Frontal lobe (“thinking, speaking, evaluating”):

- planning, deliberate behavior, imagination
- prefrontal cortex: planning, emotion judgement
- orbitofrontal cortex (OFC): situation \Leftrightarrow reward/punishment
- Broca’s area: speech production; right hemisphere: mirror neurons

Parietal lobe (“body sensing, acting, integration”):

- body sensing, action initiation, multisensorial integration;
- somatomotor cortex: action initiation; may be viewed as distinct from parietal lobe;
- somatosensory cortex: touch, pressure, vibration, tickle, itch, heat/cold, pain, body pose, movement, internal organ sensing.

Occipital lobe (“seeing, visualizing”):

- primary visual cortex area V1-V5
- from dots, edges, colors, texture, and optical flow to shapes, motion, & meaning

Brain anatomy & function, continued

Temporal lobe (“recognizing, understanding, remembering”)

- concepts recognized via signals from V5 (ventral visual pathway);
- long-term memory = episodic + semantic memory;
- face recognition (fusiform face area)
- linked to hippocampus (storage/retrieval)
- auditory cortex; Wernicke’s area (speech recog.) overlaps temporal/parietal lobes.

Cerebellum (“routine physical behavior”)

- coordination, posture, balance, motor learning, attention, language, pain;
- folding like accordion gives striated appearance

Thalamus (“grand central relay station”)

- egg-sized (& -shaped) midbrain mass (in each hemisphere)
- contains large nuclei for relaying visual, auditory, touch, etc. signals to cortices;
- smaller intralaminar nuclei (ILNs, within an internal sheet-like layer) produce the 40hz “consciousness signalling”;
- these signals are gated by the reticular nucleus, a thin sheet over the thalamus.

Corpus callosum (“hemispheric cross-traffic”)

- massive band of fibers combining brain halves;
- needed for a unique, integrated self-model (splitting the CC splits your “selves”!)

Main ideas to take from Baars ch.3

Ch. 3 pp.62-94, but see especially 62-66, 73-78

Examples of various levels of C; visual (& other sensory) functioning

pp.62-66 (consciousness of perceptions, imagined/remembered entities, thoughts):

- Faintness of mental images vs. real ones; sensory bias of C
- visio/audio/tactile traces of immediately preceding events
- forebrain : abstraction, planning, conscious control of action
- V1 - V5
- Crick & Koch

pp.73-78 (varieties of inner life):

- internally generated qualia other than imagery?
- inner voice
- self-talk, inner theater;
- phobias, hallucinations
- abstract concepts
- are ideas conscious?
- recollection of wording

Relevance to Robot Consciousness [my comments]

Access Consciousness

How conscious are we of knowledge we access when thinking or speaking?

- Consider "*The trophy won't fit in the suitcase -- it's too small*";
Why did you understand "it" as "*the suitcase*"?
Unconsciously used knowledge: *A container can't be smaller than its contents.*
- But if I ask explicitly "*Can a (closed) container be smaller than its contents?*",
the retrieved knowledge becomes conscious – it enters Baars' Global Workspace.
- Does this depend on putting the knowledge into words? Recall: "inner voice"

Would we have something analogous in an AI system with a self-model?

- This depends on what's accessible to the self-model in a thought process (e.g., QA).
- Suppose that the system is aware, and can say it is aware, of a proposition "*phi*"
if and only if "phi" is in WM (GW).
- 1. Suppose further that in processing a sentence, the system puts into WM *only that sentence, and its final internal representation*, letting the processing happen unconsciously (without putting anything else into WM).
Then the system won't be conscious of the KB facts accessed for the processing.
- 2. Suppose instead that in processing a sentence, the system puts into WM every proposition "*phi*" in its KB that it needed to access to understand the sentence.
Then, in processing "*The trophy won't fit in the suitcase -- it's too small*",
the system will come to be consciously aware that *a container can't be smaller than its contents*", and can express its awareness of this in words.

Does this sort of link between consciousness and verbalization *explain access C*?

Relevance to Robot Consciousness [continued]

Phenomenal Consciousness and the Show-and-Tell conjecture

Consider this situation:

- This is evidence of *self-awareness* (access to knowledge about oneself);
- But to *know* the bell is ringing isn't the same as *hearing* it ring! (Perhaps the man is deaf.)



What's the difference between *knowing* and *experiencing*?

A possibility, in terms of the internal self-model:

Tell only

(Person Me) etc.
(Shake Me Bell3)
(Ringing Bell3)

All constituents get their meaning *indirectly*, through their intended (and shared) correspondence to world entities

Show & Tell

(Person Me) etc.
(Shake Me Bell3)
(Ringing Bell3)

Some constituents are *directly* meaningful -- they are "signal fields" standing *for themselves*

NB: The "captioning sunglasses" would only provide the Tell version!

