

# PHILOSOPHY AND THE BRAIN

Gert-Jan C. Lokhorst

Faculty of Philosophy

Erasmus Universiteit Rotterdam

The Netherlands

University of Helsinki, 14–15 May 2001

# Contents

<b>I</b>	<b>Classical Antiquity</b>	<b>6</b>
1	Encephalocentrists vs. cardiocentrists	7
2	Early neuroanatomy	7
3	Dualists vs. materialists	10
4	Galen (2nd ct. AD)	11
<b>II</b>	<b>Middle Ages</b>	<b>18</b>

1	Ventricular localization theories	18
2	Philosophical developments	33
<b>III</b>	<b>Renaissance</b>	<b>33</b>
1	Decline and fall of ventricular localization	33
2	Ventricular localization in retrospect	37
<b>IV</b>	<b>Seventeenth/eighteenth centuries</b>	<b>38</b>
1	The scientific revolution	38

<b>2</b>	<b>Echoes of Herophilus</b>	<b>39</b>
<b>3</b>	<b>Dualisms</b>	<b>41</b>
<b>4</b>	<b>Monisms</b>	<b>46</b>
<b>V</b>	<b>Nineteenth century</b>	<b>51</b>
<b>1</b>	<b>Neuroscience</b>	<b>51</b>
<b>2</b>	<b>Philosophy</b>	<b>53</b>
<b>VI</b>	<b>Twentieth century</b>	<b>56</b>

<b>1</b>	<b>Philosophical theories</b>	<b>56</b>
<b>2</b>	<b>Mathematical models of the brain</b>	<b>70</b>
<b>3</b>	<b>The dynamical approach in cognitive science</b>	<b>78</b>
<b>4</b>	<b>Developments in neuroscience</b>	<b>85</b>
<b>5</b>	<b>Are the neurosciences fundamentally deficient?</b>	<b>89</b>
<b>VII</b>	<b>The hard problem</b>	<b>90</b>

Part I

# Classical Antiquity

# 1 Encephalocentrists vs. cardiocentrists

Alcmaeon of Croton (500 BC), Hippocrates (*On the sacred disease*), Plato (*Phaedo*, *Timaeus*) vs. Aristotle, Stoics

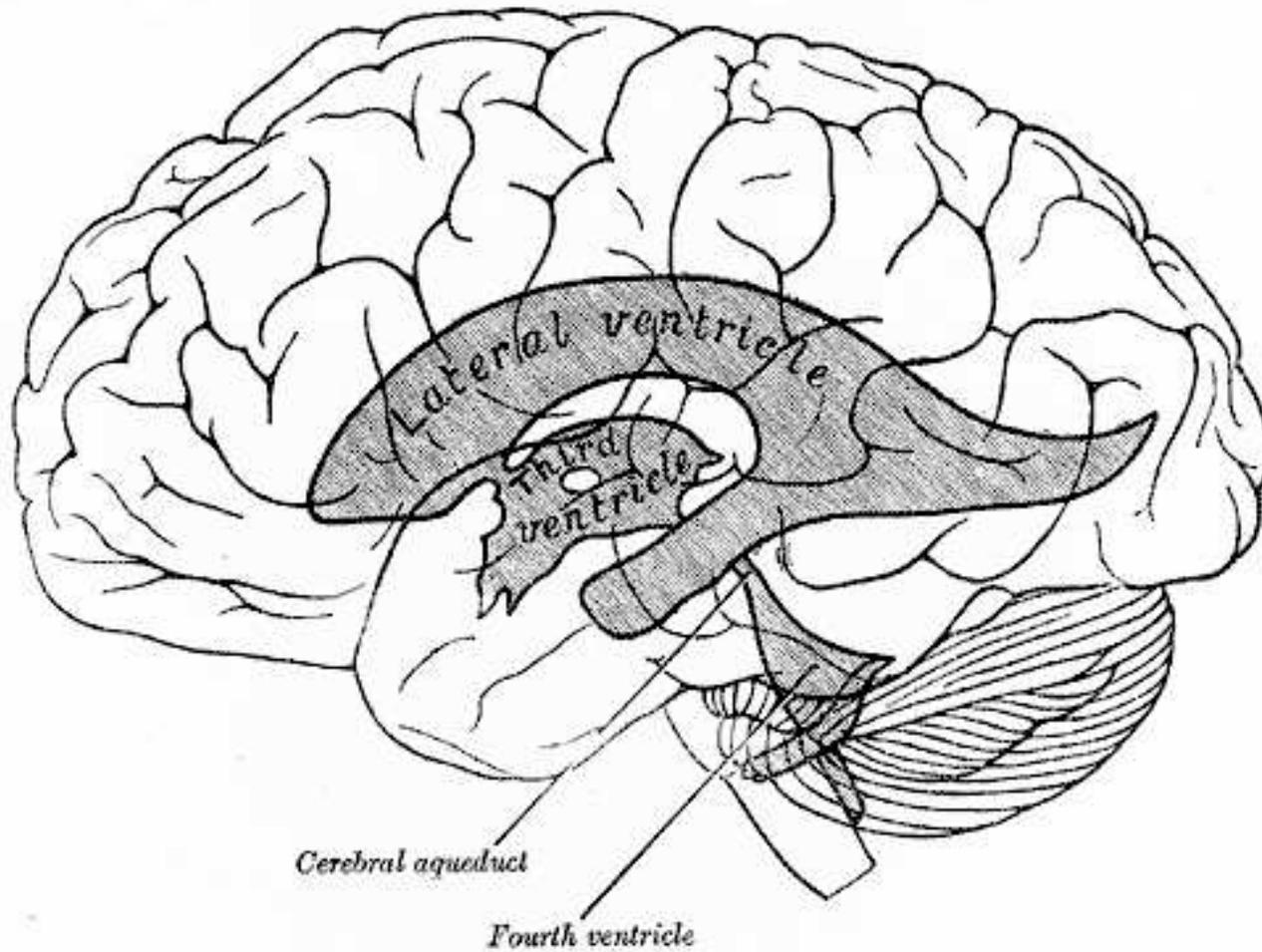
Encephalocentrism won.

# 2 Early neuroanatomy

Herophilus of Chalcedon, Alexandria, 3th ct. BC “The father of neuroanatomy.”

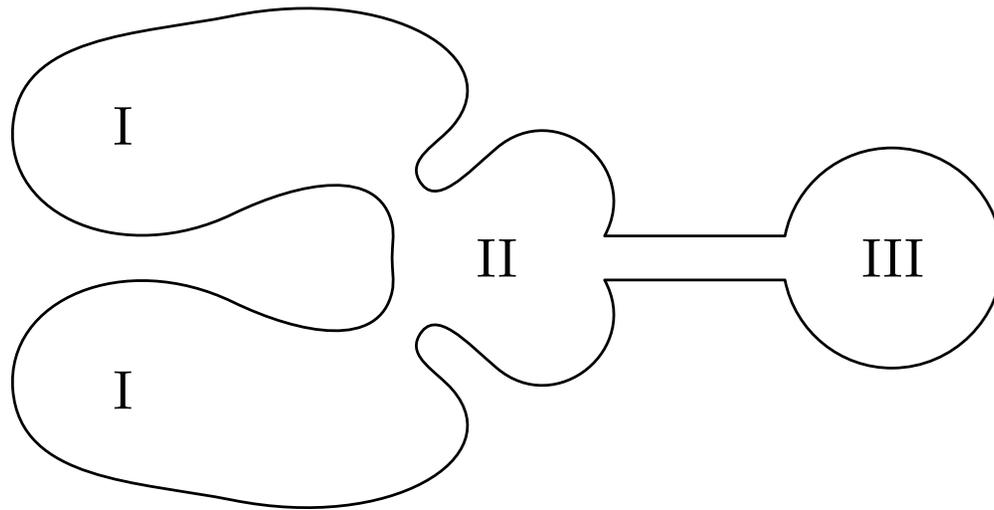
Dissection and vivisection on humans.

Discoverer of the nerves & the cerebral ventricles.



(Gray)

Herophilus regarded the posterior (fourth) ventricle as the seat of the leading part of the soul.



Terminology:

	Until Vesalius	Now
I	first or anterior ventricle (paired)	lateral ventricles
II	second or middle ventricle	third ventricle
III	third or posterior ventricle	fourth ventricle

### 3 Dualists vs. materialists

Most important dualist: Plato. Later: St. Augustine & other Church fathers.

Materialists: Stoics (soul consists of an ethereal but material *pneuma*, atomists (soul consists of a special kind of atoms), Aristotle? (soul is the “form” or “first actuality” of the living body), Tertullian (early Church father; soul consists of blood in and around the heart).

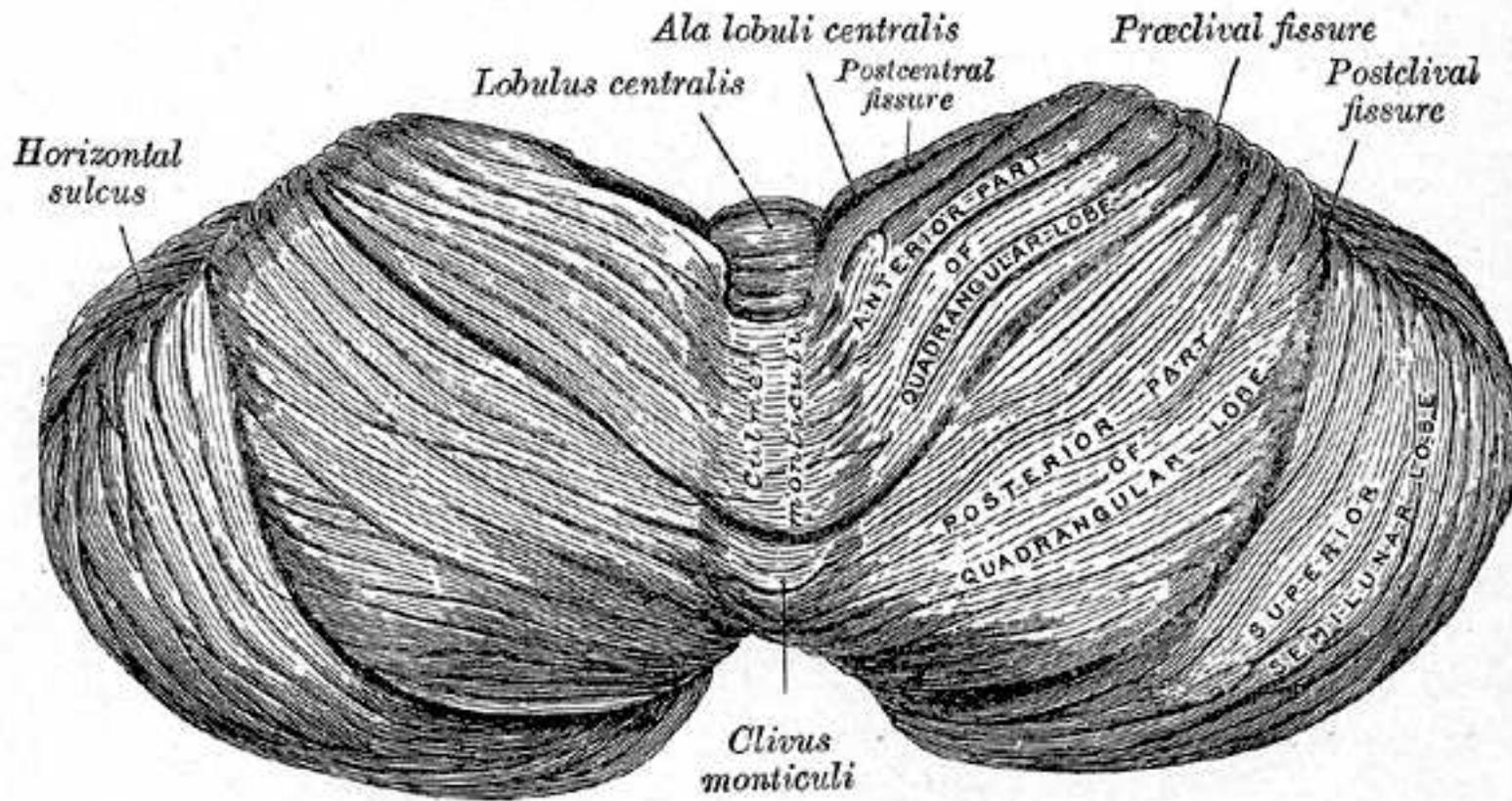
Aristotle has also been regarded as a dualist (e.g., by Averroës).

## 4 Galen (2nd ct. AD)

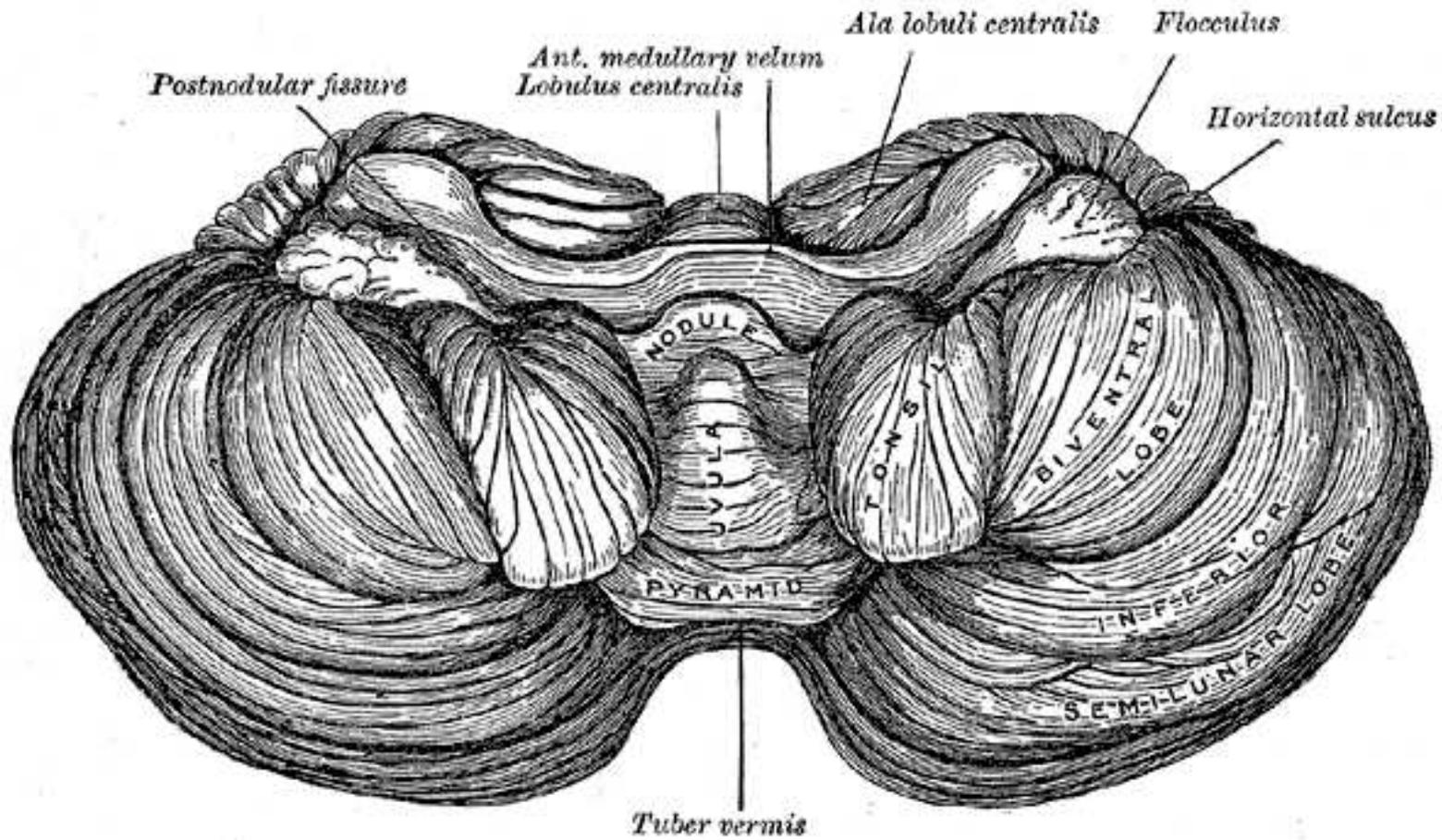
Two kinds of *pneuma*:

- vital spirit (heart, blood vessels)
- animal spirit/psychic pneuma (*rete mirabile* and ventricles of the brain)

The *vermis* (“worm”) of the cerebellum regulates the flow of animal spirit; the pineal gland does *not* have this function.



Upper surface of the cerebellum (Gray).



Under surface of the cerebellum (Gray).

The animal spirit is the *primary instrument* of the leading part of the soul, which is associated with the substance of the cerebrum. G. did not say whether he regarded it as material or immaterial.

Localization of the most important functions of the soul (*On the differentiation of symptoms*, ch. 3):

- sensory—anterior, soft part of the brain (cerebrum)
- governing:
  - *phantasia*
  - reasoning
  - memory & recollection
- motor—posterior, hard part of the brain (cerebellum)

*Phantasia*, reasoning and memory may become disturbed independently from one another:

- case 1 (disturbed *phantasia*: imaginary flute players),
- case 2 (disturbed reason: man throwing objects out of a window)
- case 3 (loss of memory, from Thucidydes)

G. regarded the middle ventricle as no less important than the posterior ventricle (*On the doctrines of Hippocrates and Plato* III.8); lesions of the middle ventricle may cause loss of reason (*On the affected places* IV.3).

G. did not put forward a full-blown ventricular localization theory.

*Galien n'a jamais assigné de sieges particuliers aux facultez princesses* (André du Laurens, *Œuvres* (Paris, 1639), p. 552)

*Lokalisationen sind allerdings keine zu finden; diese erscheinen erst viel später* (W. Sudhoff, *Sudhoffs Archiv* VII (1913), p. 156)

## Part II

# Middle Ages

## 1 Ventricular localization theories

### Early theories

Porphyry (233–304)

Lost commentary on Plato's *Timaeus*.

**Posidonius of Byzantium (4th ct.)**

(Aëtius of Amida, *Medical books*, VI.2.)

<i>phantasia</i>		anterior part of the brain
reason		middle ventricle
memory		posterior part of the brain

**Bishop Nemesius of Emesa (ca. 400)**

In his book *On the nature of man*.

<i>phantasia</i>	anterior <u>ventricle</u>	}	sensory nerves
reason	middle ventricle		
memory	posterior <u>ventricle</u>		motor nerves

**Saint Augustine (354–430)**

In his book *Genesis to the letter*.

perception	anterior ventricle
<u>memory</u>	middle ventricle
motion	posterior ventricle

The theory of Nemesius was the standard model until Avicenna (11th ct.). Examples:

- St. John of Damascus (8th ct.)
- Costa ben Luca (†923)
- Rhazes (†923)
- Ibn al-Jazzār (†980)
- Haly Abbas (†994)

Rhazes (†923)

First depiction of the ventricles.

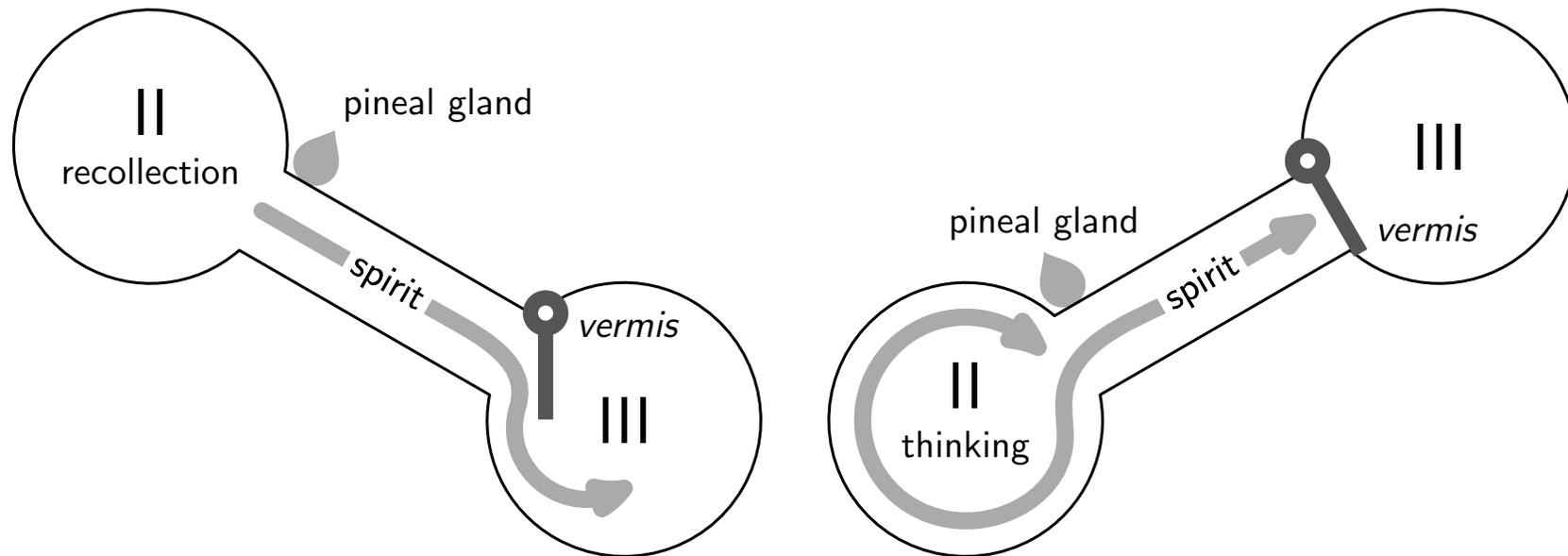
وفي ظهر القدمين بالقرب من الوتر العظيمة  
السابع في الدماغ

ان الدماغ ليس بمصمت لكن له تجاوير وهي على راي جالينوس اربعة  
تجاوير بعضها التي تسمى بطون الدماغ اثنان منها في مقدم الدماغ وواحدة  
في وسطه واخر في مؤخره على هذا الشكل ° ° ° ° وعند هذه الحمار  
اجام مشكلة بكل موافق يدها في بعض الاحايين ويفتحها في اخرى وله

بعض صور

## Costa ben Luca (†923)

*On the difference between the spirit and the soul. The role of the vermis:*

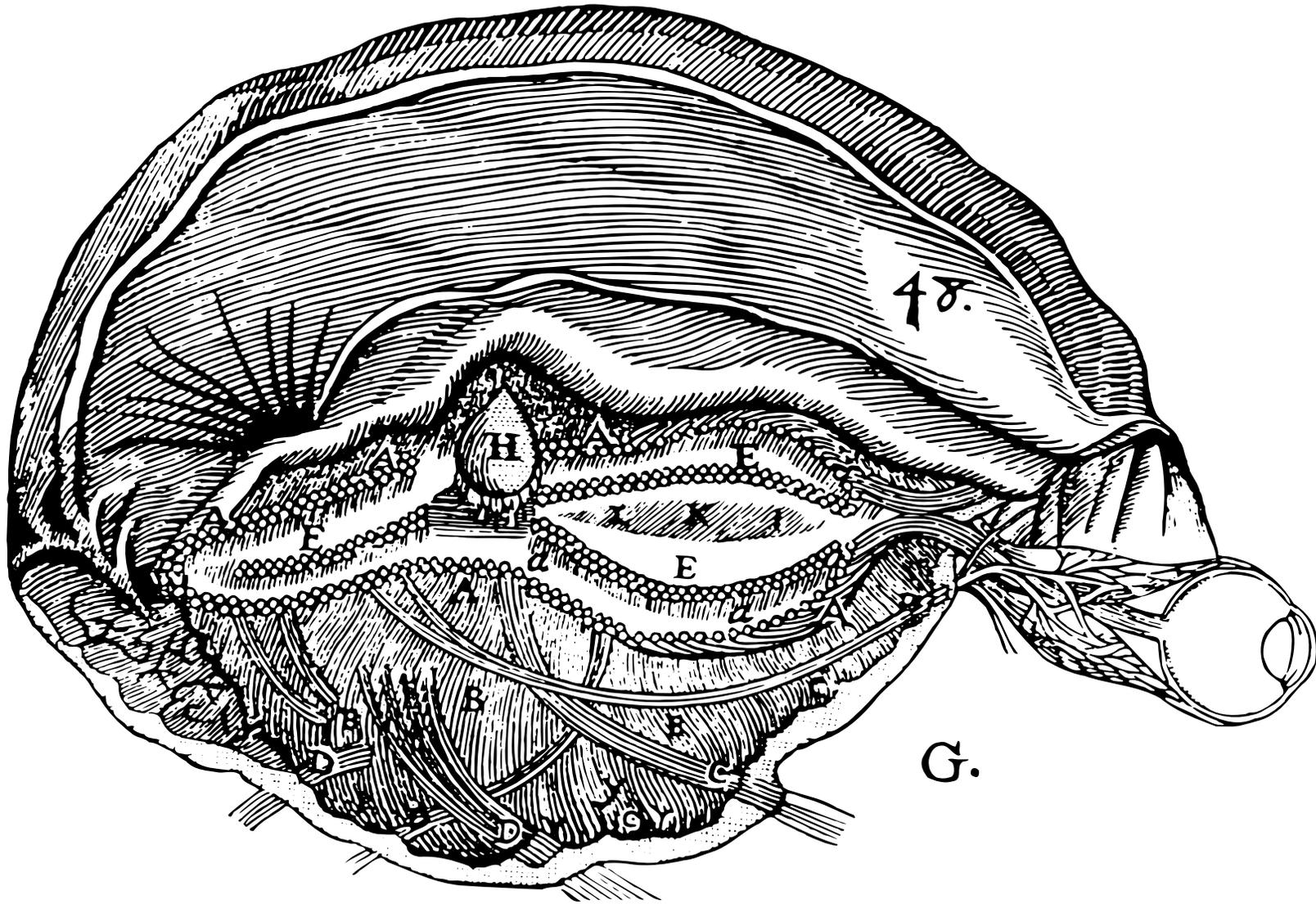


## Intermezzo: the *vermis* and the pineal gland

The *vermis* was sometimes confused with the pineal gland, with the result that this organ came to be regarded as the governor of the flow of animal spirit (an idea that had been rejected by Galen):

- Ibn al-Jazzār (†980), *On Forgetfulness and its Treatment*,
- Constantine the African (†1087?), *Liber de oblivione* (11th ct., printed in 1515),
- Vincent of Beauvais, *Speculum naturale* (13th ct., printed in 1624), book 27, ch. 12.

This idea recurs in the works of René Descartes (1596–1650).



## Later developments

### Avicenna (980?–1037)

There are *five* “internal senses” (“inward wits”)

“fantasia” (sensus communis)	I (frontside)
imaginatio (formans)	I (backside)
imaginativa & cogitans <sup>a</sup>	II (frontside)
æstimativa	II (backside)
memorialis & reminiscibilis	III

---

<sup>a</sup>“cogitans” only in humans

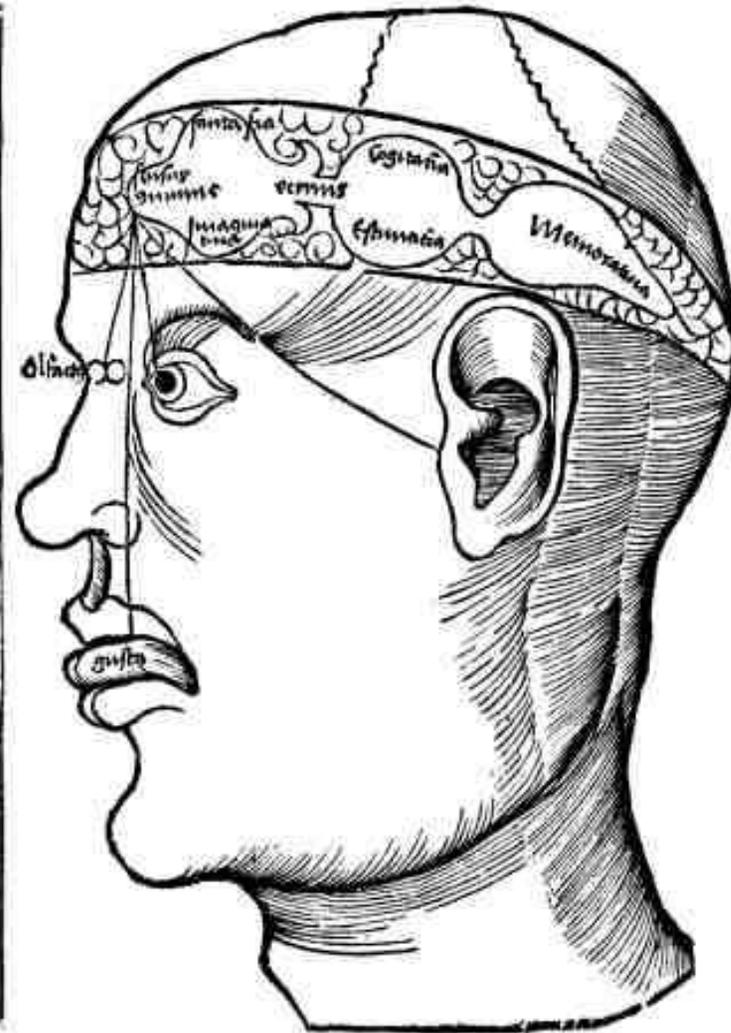
$$\overbrace{\text{fantasia : imaginatio}}^{\text{formæ}} \approx \overbrace{\text{æstimativa : memorialis}}^{\text{intentiones}}$$

This was the standard model until Vesalius (1543).

Avicenna located the *vermis* in the anterior part of the middle ventricle (between *imaginatio* and *æstimativa*) and said that it regulates the flow of psychic pneuma (*spiritus animalis*) by contracting and extending like a worm.



The posterior worm (Galen)



The anterior worm (Avicenna)

## Variants

Albert the Great (1193?–1280), *De Anima*:

Ia	Ib	IIa	IIb	III
sensus communis	imaginatio	<u>æstimatoria</u>	<u>“phantasia”</u> (imaginativa)	memoria

Averroës (1126–1198), Thomas Aquinas (1225–1274):

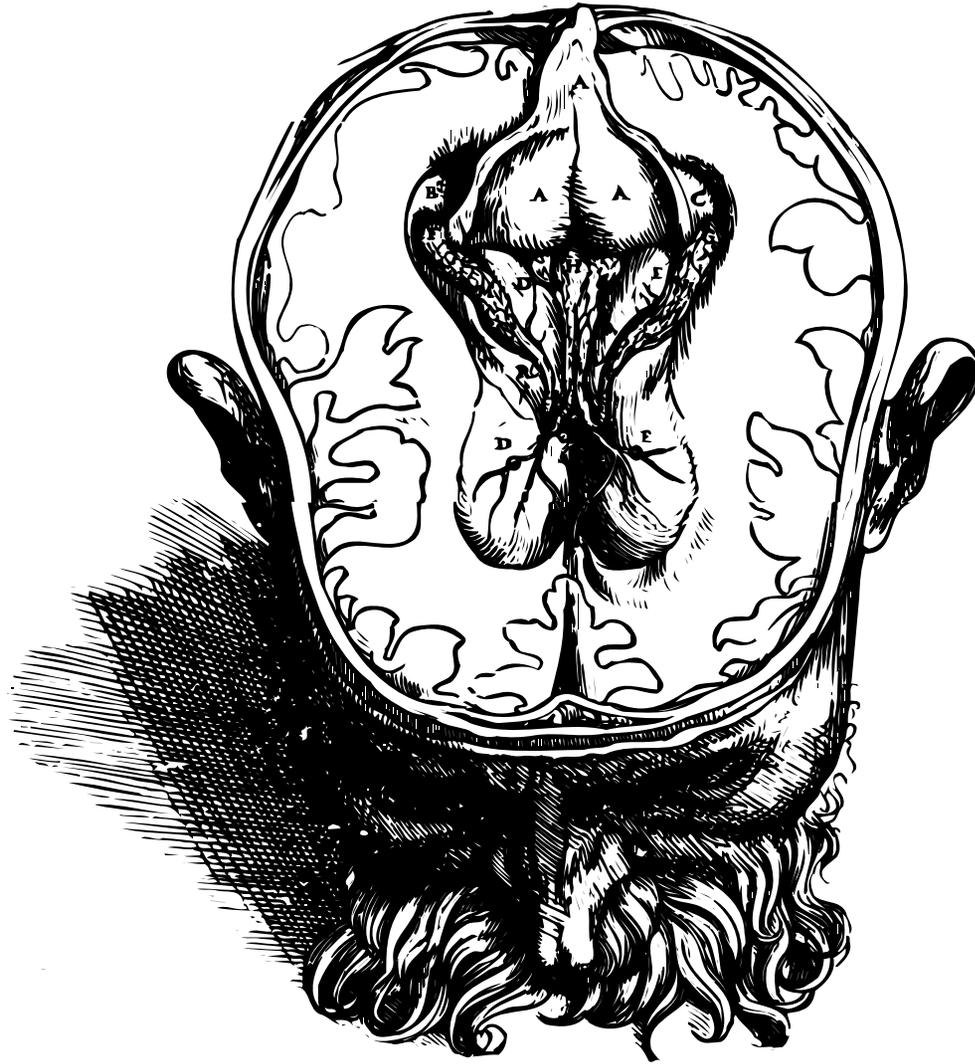
Ia	Ib	II	III
sensus communis	imaginatio	æstimatoria ( <i>animals</i> ) cogitativa ( <i>humans</i> )	memoria

(no *imaginativa*)

Mondino dei Luzzi (1270?–1326), *Anothomia* (1316):

Ia	Ib	Ic	II	III
<u>“phantasia”</u> (imaginatio)	<u>sensus communis</u>	<u>imaginativa</u>	cogitativa	memoria

Mondino identified Avicenna’s *vermis* with the choroid plexuses of the lateral and third ventricles.



(Vesalius)

## 2 Philosophical developments

All medieval authors we have mentioned were dualists; they saw the soul as immaterial and immortal. But they did not go into the details. E.g., Avicenna: the *æstimatoria* “shows” (*ostendit*) the sensible forms to the soul.

There were two main problems: (1) reconciling Aristotle (cardiocentrist) with Galen (encephalocentrist); (2) reconciling Aristotle (the soul is the form of the body and mortal) with Plato (the soul is separable from the body and immortal); hence the doctrine of “substantial forms.”

## Part III

# Renaissance

## 1 Decline and fall of ventricular localization

Beginning of the 16th ct.:

- Galen's works became better known and much more influential; "the Arabs" became unpopular.
- Anatomical research showed that the ventricles are filled with liquid (liquor cerebrospinalis) rather than air/pneuma/spirit.
- Vesalius (1514–1564), *De humani corporis fabrica* (1543).

Vesalius ridiculed the ventricle theory proposed by “Thomas, Scotus, Albertus, and theologians of that school” (book VII, ch. 1).

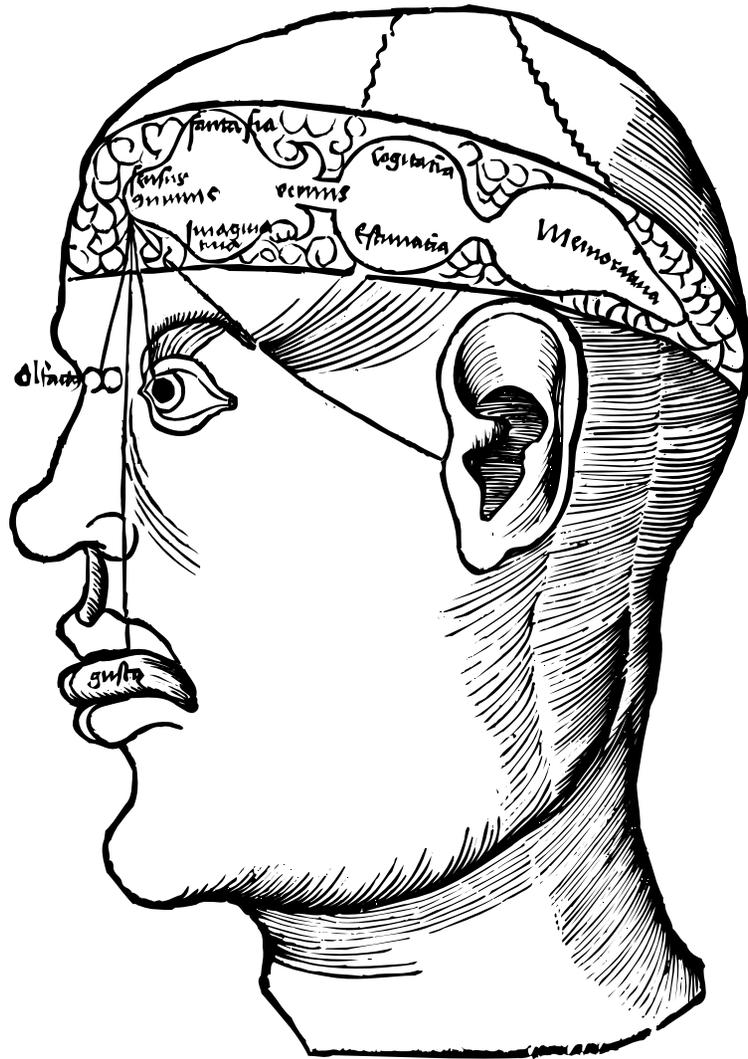
His main argument:

All our contemporaries, so far as I can understand them, deny to apes, dogs, horses, sheep, cattle, and other animals, the main powers of the Reigning Soul—not to speak of other [powers]—and attribute to man alone the faculty of reasoning; and ascribe this faculty in equal

degree to all men. And yet we clearly see in dissecting that men do not excel those animals by [possessing] any special cavity [in the brain]. Not only is the number [of ventricles] the same, but also all the other things [in the brain] are similar.

Vesalius explicitly referred to the following diagram from Gregor Reisch’s *Margarita philosophica* (1503).

**De potentijs anime sensitive**



G. Reisch, *Margarita philosophica*, Basileæ 1517, Liber X Tractatus II Capitulum XXI, fol. E vij recto.

I { fantasia  
sensus communis  
imaginativa

vermis

II { cogitativa  
estimativa

III memorativa

[= Mondino + estimativa]

Vesalius denied that the *vermis* (either Galen's *vermis* or the Avicennian *vermis*, i.e., the choroid plexus) regulates the flow of animal spirit.

## 2 Ventricular localization in retrospect

Nigel J. T. Thomas, "Imagination, Eliminativism, and the Pre-History of Consciousness," Paper delivered at the *Toward a Science of Consciousness* (Tucson III) Conference, April 30th 1998.

We can easily abstract away from the lousy neuroscience. We are then left with a cognitive theory that is not so very different from the sort of thing that you might find in contemporary textbooks of

Cognitive Psychology (figure 2) ... Contemporary cognitive theory labels its boxes differently, but their functions do not seem so very different.

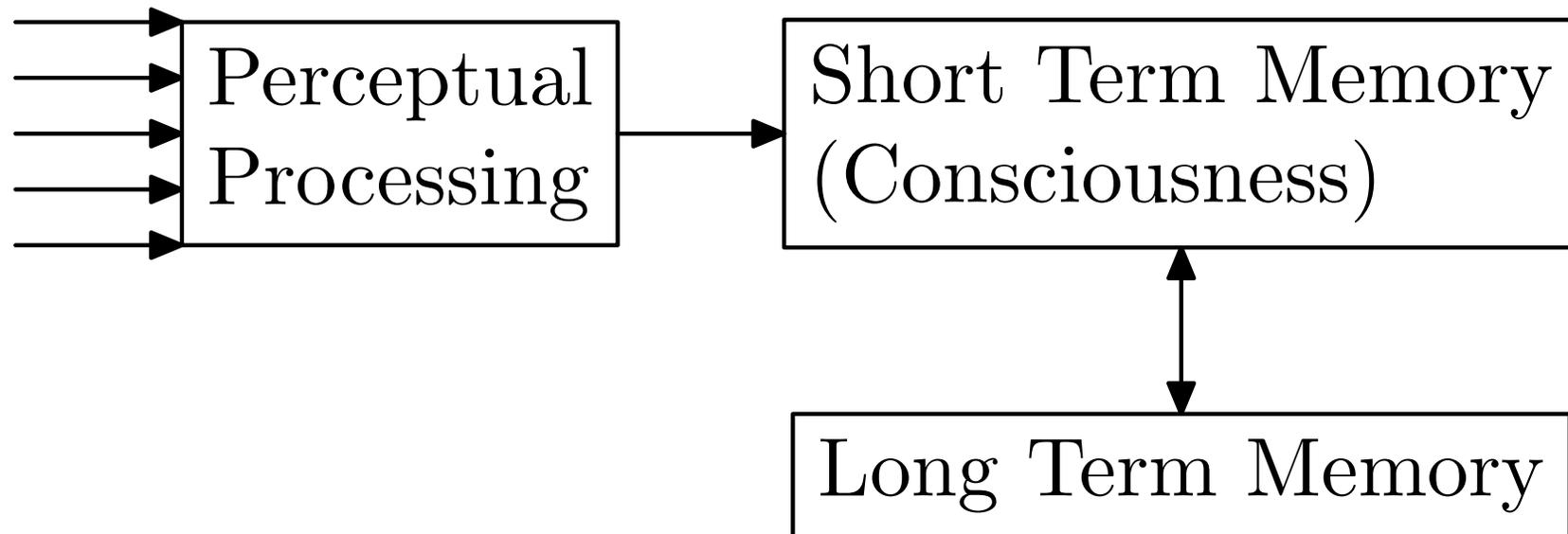


Figure 2. The bare bones version of contemporary Cognitive theory.

## Part IV

# Seventeenth/eighteenth centuries

## 1 The scientific revolution

Rejection of Aristotle (Galileo Galilei), deanimation of matter. This raised a problem: what is the place of the rational human soul in the natural world?

## 2 Echoes of Herophilus

Henry More (1614–1687): the posterior ventricle is the seat of the *sensus communis* (*An Antidote against Atheism*, 1653, *The Immortality of the Soul*, 1659)

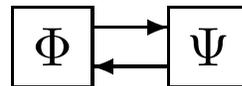
Thomas Sömmering (1755–1830): the liquor in the ventricles is the seat of the *sensus communis* or soul (criticized by Immanuel Kant, 1796).

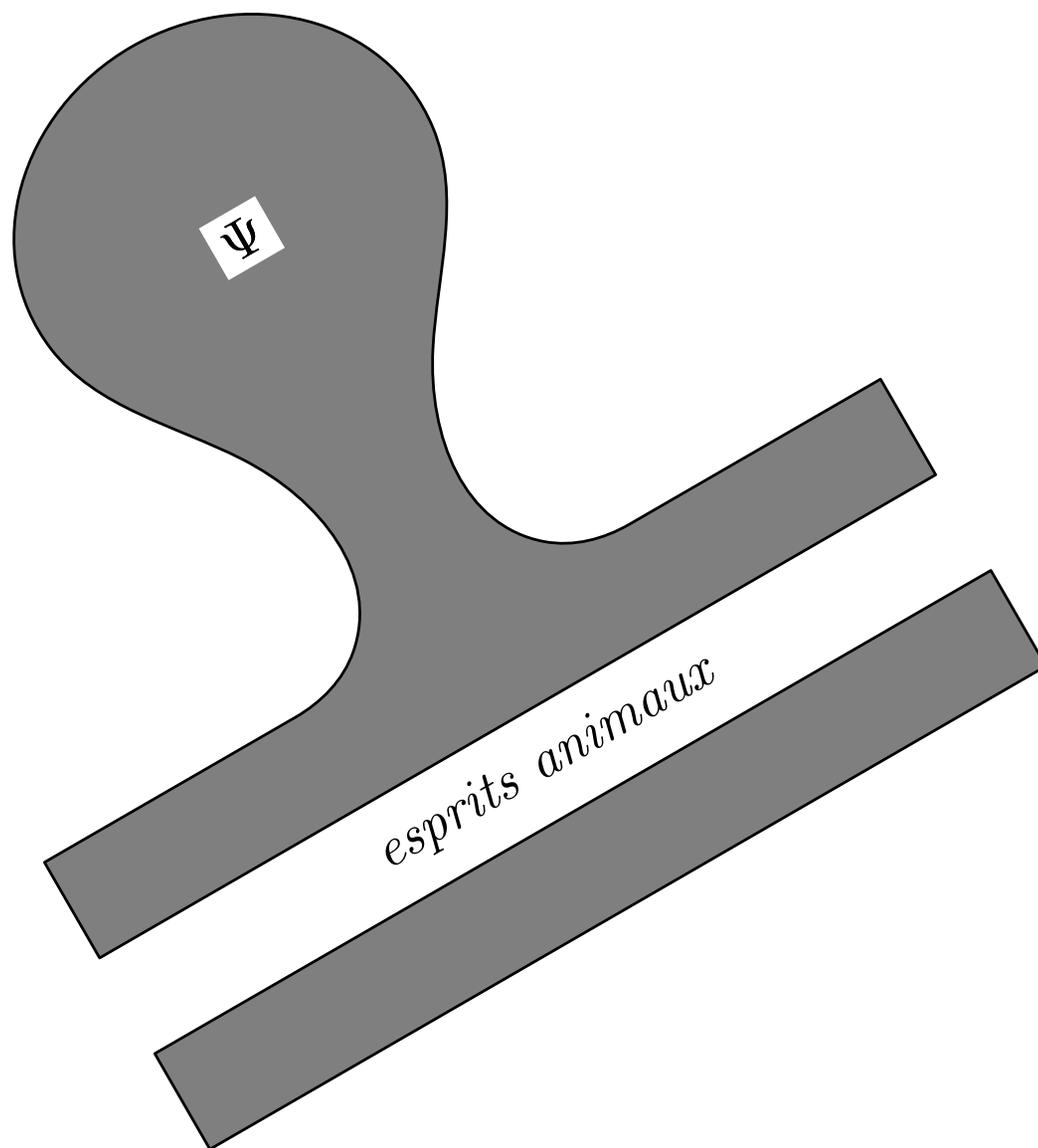
But in general, attention shifted to the *substance* of the brain (cf. Galen); first, the pineal gland (Descartes); later, other structures of the brain (Willis a.o.); in the eighteenth century, the *corpus callosum* was widely regarded as the seat of the soul (Lancisi, Diderot's *Encyclopédie*). This hypothesis was refuted by Zinn (split-brain experiments on dogs).

### 3 Dualisms

#### Cartesian interactionism

Two substances (bearers of properties): the soul (anima, l'âme) (*res cogitans, non extensa*) and matter (*res extensa, non cogitans*), which are in causal interaction in the pineal gland





### Problems:

- Theory does not explain anything. Why is the soul able to think, perceive and so on, and is the body unable to carry out these activities? What would a testable theory of, e.g., the soul's perceptual abilities look like?
- Soul-body interaction is *incomprehensible*.
- Violation of physical conservation laws.

## Occasionalism (Malebranche, Geulincx a.o.)

There is no interaction. God takes care of synchronisation.



## Parallellism (Leibniz)

No interaction. But God does not intervene on every occasion. There is a *harmonia præstabilita*.



[But: How can we be sure that mind and body operate synchronously?—Smullyan]

## 4 Monisms

### Spinoza (1632–1677)

There is only substance, *deus sive natura*, which has many kinds of attributes (aspects); only two of these are known to man: *cogitatio* and *extensio*.

Problem: why postulate an intrinsically unknowable *tertium quid*?

## Materialism

End 17th century (medical circles in the UK), 18th century (France).

Most prominent materialist: Lamettrie (1709–1751), author of *Traité de l'âme* (1745), *L'homme machine* (1748), *L'homme plante* (1748).

The excellence of reason does not depend on a big word devoid of meaning (immateriality), but on the force, extent, and perspicuity of reason itself. Thus a “soul of clay” which should discover, at one glance, as it were, the relations and

the consequences of an infinite number of ideas hard to understand, would evidently be preferable to a foolish and stupid soul, even if that were composed of the most precious elements.

Galen emphasized that “the power of the soul follows the complexion of the body” but Lamettrie went further: he seems to have thought that all mental processes are *completely determined* by physical processes (thesis of psychophysical supervenience, to be discussed later).

Well-known twentieth-century books with essentially the same message:

- Jean-Pierre Changeux, *L'homme neuronal* (1983).
- Francis Crick, *The Astonishing Hypothesis: The Scientific Search for the Soul* (1994).

The Astonishing Hypothesis is that “You,” your joys and your sorrows, your memories and your ambitions, your sense of personal identity and free will, are in fact no more than the behavior of a vast assembly of nerve cells and their associated molecules. As Lewis Carroll’s Alice might have phrased it: “You are nothing but a pack of neurons.” This hypothesis is so alien to the ideas of most people alive today that it can truly be called astonishing.

Lamettrie compared man to a clockwork; in our times, the digital computer is the favourite object of comparison. We will return to this point later on.

## Part V

# Nineteenth century

## 1 Neuroscience

### Franz Joseph Gall (ca. 1800)

- The cerebral *cortex* is the organ of the mind
- Localization of function ( $\implies$  phrenology)

## Broca

Speech center (1861)

Cerebral dominance (1863)

## Others

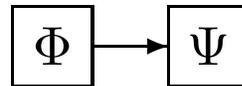
Discovery of neurons, finite speed of signal propagation in neurons (Helmholtz), etc.

## 2 Philosophy

### Epiphenomenalism

(Hodgson, Thomas Huxley)

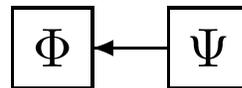
The mental is an epiphenomenon (by-product) of neurophysiological activity.



Problem: please specify the details.

## “Reverse epiphenomenalism”

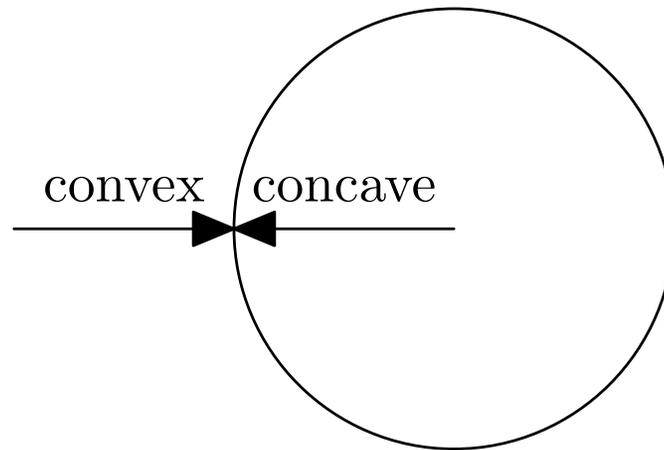
Christian Science (Mary Baker Eddy)



Problem: ludicrous. Think of Alzheimer's disease and so on.

## Dual aspect theories

Gustav Fechner, Lewes. Two aspects or perspectives.



Objection: sounds rather mystical.

## Part VI

# Twentieth century

## 1 Philosophical theories

### Epistemological neutral monism

- Mach, James, Russell. The world consists of neutral ‘events’, which are classified in two distinct ways. Physical and mental phenomena are “logical constructions.”

Assessment: OK from a first person point of view. Less OK from a third person point of view (“other minds” problem; many questions about the relationship between mind and body remain unanswered, e.g., is the thesis of psychophysical supervenience tenable or is it not?).

Moreover, logical reconstructionism is generally considered as a failed research program.

## Physical monism, physicalism

The natural sciences have the last word about what exists and what does not exist. Quarks, electromagnetic fields, and so on, seem to exist, but there are no psions and there is no element  ${}^0_0\text{Psychium}$  in the periodical table of the elements.

## Emergentism

Already to be found in Lamettrie. Sufficiently complex well-organized structures give rise to minds. Weak vs. strong emergence. Weak emergentism (a whole may have properties that none of its proper parts has) is uncontroversial, strong emergentism (these emergent properties may have their own “downward” causal effects) is incomprehensible.

## Contingent token-token identity theories

(fifties and sixties).

Each occurrence of a mental event is contingently identical with an occurrence of a physical event in the central nervous system.

Similarly, mental processes are physical processes and mental states are physical states.

- Why ‘contingent’?
- What is ‘token-token’? Examples: picknicks, thinking about the sinking of the Titanic.

Cf. Davidson’s “anomalous monism”.

## Psycho-physical supervenience theories

(eighties).

The mental is wholly determined by, supervenes on the physical.

No mental difference without some physical difference.

**Formal explication of the supervenience thesis** There are two sorts of properties,  $B$  (base properties) and  $S$  (supervenient properties).  $b$  ranges over  $B$ ,  $s$  over  $S$ . The supervenience thesis is:

$$(1) \quad \Box \forall x \forall y (\forall b (bx \supset by) \supset \forall s (sx \supset sy))$$

$B$  should not be taken too narrowly and  $S$  should not be taken too widely. For example, it is clearly false to say that propositional attitudes supervene on properties of the central nervous system.

Veridical propositional attitudes and *de re* propositional attitudes do

not supervene on properties of the central nervous system. Examples:

- “We know that Bush is the president of the USA.” (This implies that Bush is the president of the USA. But he may just have died. Thus, our knowledge may just have turned into belief, without our nervous systems having anything to do with the matter.)
- “We believe of the president of the USA that he is called Bush.” (This implies that there is a president of the USA, which may just have ceased to be the case. Thus, the propositional attitude ascription in question may just have become false while our nervous systems have remained the same.)

It has been claimed that supervenience does not entail reducibility, but this claim should be taken with a grain of salt (as has been shown by John Bacon).

(1) is equivalent with (2):

$$(2) \quad \Box \forall s \forall x (sx \equiv \bigvee \{ \bigwedge \{ b : by \} : sy \} x).$$

If  $B$  is closed under Boolean operations, then (2) is equivalent with (3):

$$(3) \quad \Box \forall s \exists b \forall x (sx \equiv bx)$$

Moreover, if  $B$  is closed under resplicing, in the sense that  $\{ f : \forall w \exists b E(f, w) = E(b, w) \} \subseteq B$  (where  $E$  is the extension of  $f$  at  $w$ ), then (3) entails (4):

$$(4) \quad \forall s \exists b \Box \forall x (sx \equiv bx)$$

But (4) is nothing less than a “reducibility-in-principle” thesis.

Difference between token-token identity theory and supervenience theory: identity theories talk about sorts of objects/events; supervenientists talk about relationships between properties of different sorts. You can be an old-fashioned dualist and a supervenientist at the same time.

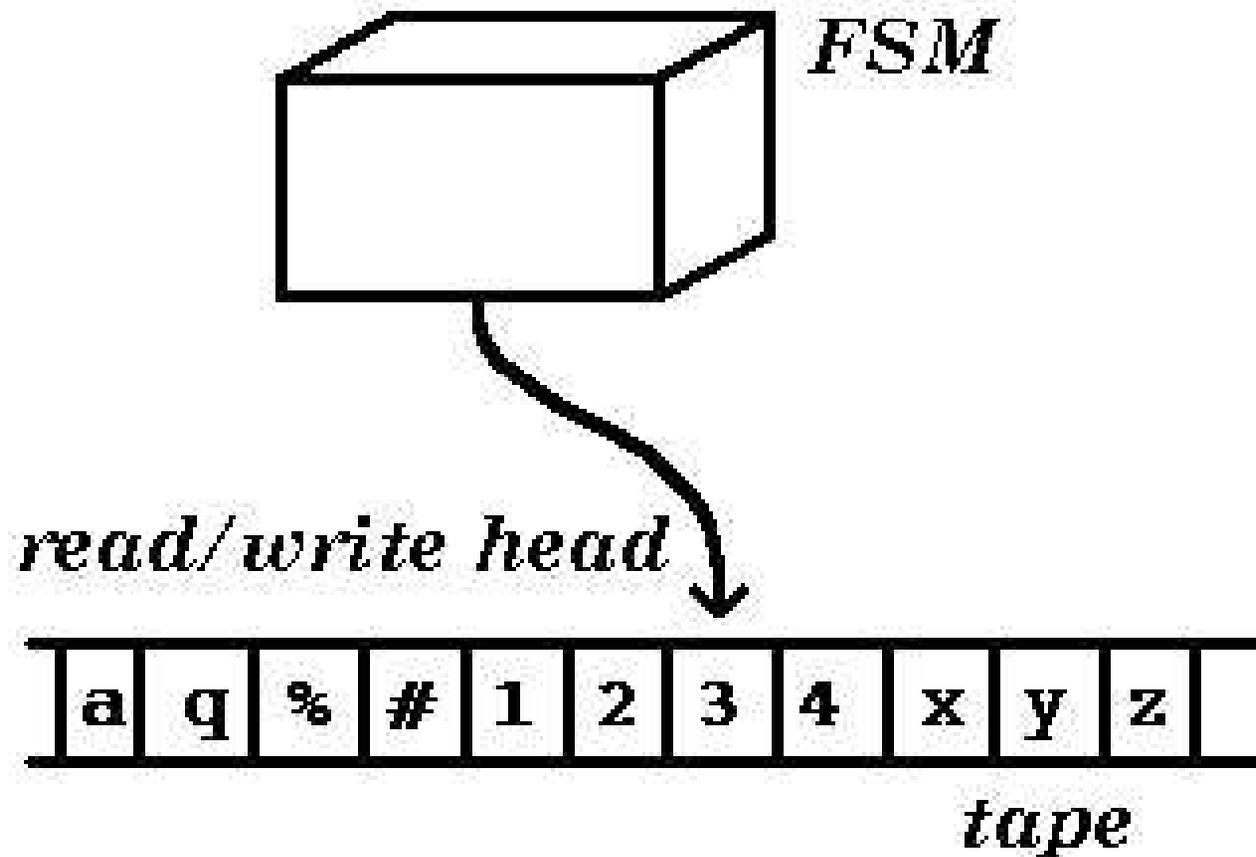
On the other hand, any physicalist must probably accept some sort of ‘global’ supervenience thesis, for otherwise the physicalist view of the world would be incomplete (if there are properties which do not supervene on physical properties, physicalism is incomplete).

Objections to these two modern theories: none [?].

**A theory which seems to have fallen out of favour:  
functionalism**

**1. Turing Machine functionalism (Hilary Putnam, 60s)**

# A Turing Machine



Mental states are somehow comparable to (“have the same logical status as”) the configurations of Turing machines.

Objections (Putnam himself):

- problem of infinite disjunctions
- ascription of mental state is a social affair
- meanings are not “in the head”
- no guarantee whatsoever that reality is computable, that the brain is a finite automaton or Turing machine

## 2. Contemporary functionalism

Based on three ideas (Rey)

- TM functionalism—see above.
- Ramsey reducibility—objection: depends on strict distinction between theoretical terms and observation terms, and on truth of the reduced theory (questionable: see eliminativism below).
- Multiple realizability—rather meager idea and perhaps questionable.

Objections to functionalism.

- Searle's Chinese room—Not OK: applicable to Searle himself.
- “Each rock implements every finite automaton” (Putnam, Searle).

This argument has been conclusively refuted (O'Rourke, Copeland, Chalmers). My computer is running under the Finnish operating system Linux, not under some lousy commercial OS. Searle has not “reinterpreted” his old home-computer as a brand-new 1500 MHz Pentium.

Associated ideas:

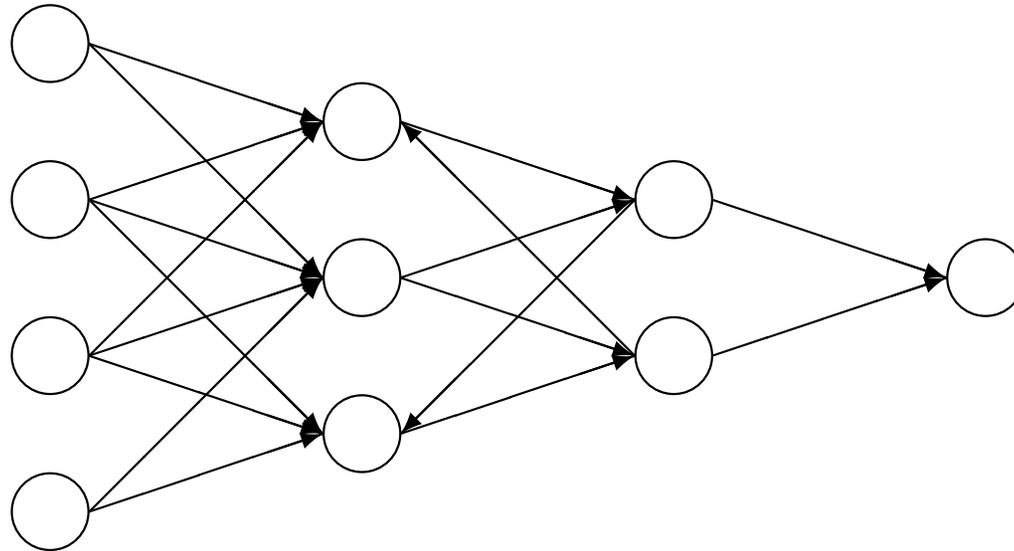
- “language of thought hypothesis” (fruitless).
- “computationalism” (too general to be interesting).

## **2 Mathematical models of the brain**

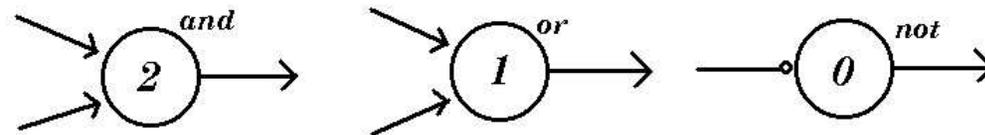
### **Neural networks**

Important source of inspiration for modern philosophy of mind.

A simple example:

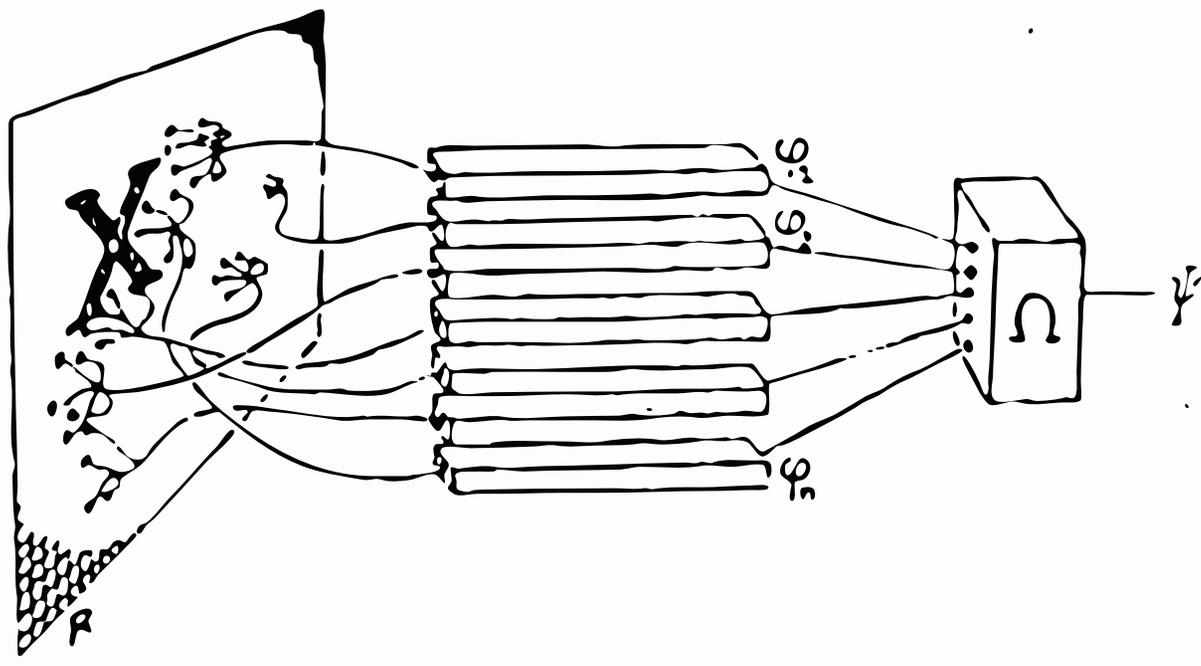


- First neural nets: McCulloch & Pitts 1943. These consisted of simple “logical neurons”:



- First one who attempted to make neural nets tolerant to noise: John von Neumann (50s).
- First one who tried to invent learning algorithms for neural nets: Alan Turing (around 1950). He did not pursue his ideas, and went on to create the field of Artificial Life.
- First one who succeeded: Marvin Minsky, PhD thesis of 1954.

- Frank Rosenblatt: the perceptron (feedforward neural network with two layers of units), with “perceptron learning rule”, inspired by Hebb



- Neural nets became very popular in the 1980s after the invention of the backpropagation learning algorithm for multi-layered feedforward networks. Most famous example: NetTalk.

Detailed description of NetTalk & attractive features of neural nets

## Computational powers of analog recurrent neural networks

(Hava Siegelmann, PhD thesis 1993)

nets with integer weights	finite automata (Kleene 1956)
nets with rational weights	Turing machines
nets with real weights	Turing machines with oracles (*)

(\*) First studied by Turing in his PhD thesis of 1939.

The neurophysiologically most adequate models of brain activity which exist today are the so-called “third generation” neural network models with **spiking neurons**. These networks are equivalent with analog recurrent networks with **real weights**, which are in turn equivalent with Turing machines with **oracles**. The latter machines are strictly more powerful than ordinary Turing machines.

This raises the following question: is man a **super-Turing machine**?

Probably not. There is one factor that spoils the fun: noise.

Spiking neural nets with noise are, in general, less powerful than finite automata. (Wolfgang Maass)

## Technological excursion

The brain contains about  $10^{11}$  neurons. There are about  $10^{14}$  connections. Suppose there are about 100 operations/second per connection, i.e.,  $10^{16}$  operations/second in total (ten petaops).

Extrapolating technological developments from the last three decades, an affordable PC with this speed will be on the market in about 50 years (Moravec)!

Human memory poses no problem at all: a person stores no more than a few hundreds of megabytes during his or her life (Landauer, Merkle). That easily fits on one CD-ROM.

How should we program such a PC to give it human intelligence? Perhaps in the same way that we program humans: by sending it to school (Turing 1948).

### 3 The dynamical approach in cognitive science

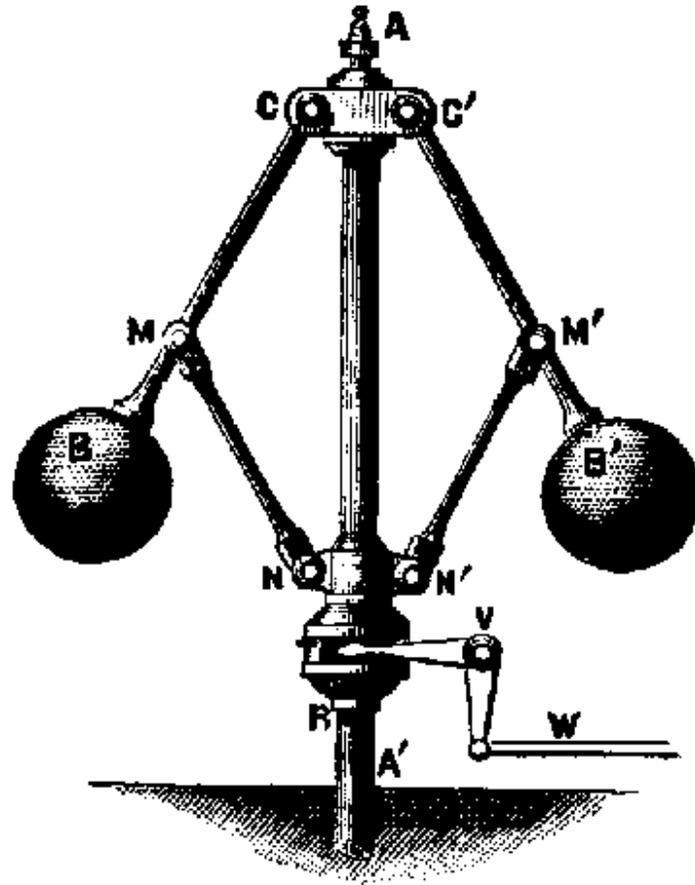
Main manifestos:

- Tim van Gelder: “What might cognition be, if not computation?” *The Journal of Philosophy* 91 (1995): 345-381.
- Tim van Gelder: “The dynamical hypothesis in cognitive science.” *The Behavioral and Brain Sciences* 21 (1998): 615–628.

Main thesis: Natural cognitive systems are to be viewed as *dynamical systems*.

Examples of dynamical systems: the solar system, the weather, James Watt’s centrifugal governor (the dynamicists’ favorite example).

# The centrifugal governor (James Watt, 1788)



Like other dynamical systems, cognitive systems are to be described in terms such as:

- state space,
- trajectories through state space,
- attractors and repellers, limit cycles,
- stability, deterministic chaos,

and so on (see any book on chaos theory).

**Objection** This is not very new or controversial. Cf. the cybernetics movement of the 50s and 60s, e.g., Ashby's *Design for a Brain* (1952). Moreover, connectionists such as Sejnowski and Churchland have employed the same terminological framework.

Dynamicism has **two specific tenets**:

- Non-computationalism
- Non-representationalism

The first term seems a misnomer. The dynamicists do not deny that their systems are computational in a broad sense of the word (even though some of them are perhaps not Turing computable). The dynamicists should have used some term like “non-digitalism”.

The second issue is more interesting.

Non-representationalists abhor talk of “internal representations”. Instead, they say that each cognitive agent is “dynamically coupled” to the environment in which it is “embedded”. There is a constant mutual interaction between the agent and its environment; they form one integrated system and cannot very well be isolated from each other. Cognition is most fruitfully seen as adaptive activity in an appropriate environment.

There are obvious points of similarity here with the philosophical positions adopted by Heidegger, Merleau-Ponty and the later Wittgenstein, a tradition which is nowadays being continued by thinkers such as Dreyfus & Freeman and Maturana & Varela. The dynamicists are well aware of this connection. They view their research programme as giving a scientific twist to the earlier, purely philosophical tradition.

On the other hand, there is a break with connectionism, which continually talks about distributed representations.

Personally, I think that the anti-representationalists are overstating their case.

- First, many cognitive capacities seem to have nothing to do with being in a tightly coupled relationship with the environment. I can reason about which retirement fund would be best in any non-distracting situation. Or think of mental imagery, doing mathematics and planning ahead: all these activities presuppose a considerable amount of *decoupling* from the actual environment. One might say that it is precisely our ability to *break out* of the feedback loop with the environment that makes us human (Popper: we let our theories die in our stead).

- Secondly, representationalism and dynamicism seem quite compatible. As Van Gelder himself put it, representations may be viewed as “trajectories or attractors of various kinds, or even such exotica as transformations of attractor arrangements as a system’s control parameters change.” As soon as representations are not naïvely identified with pictures or propositions but with some more sophisticated constructs, there is ample room for them in the dynamical framework.

[Notice that even the centrifugal governor seems to represent: the height of the balls represents the engine’s speed.]

## **4 Developments in neuroscience**

Consciousness is no longer a taboo (Francis Crick).

Example of recent work: visual awareness.

- Crick & Koch 1994: visual awareness is due to activity in the pyramidal cells of layer 5 of visual area V1.
- Crick & Koch 1995: no, this cannot be true, because this area is not directly connected with the prefrontal cortex, which gives rise to motor activity.
- Ned Block (a philosopher!) 1996: one should make a distinction between *phenomenal consciousness* and *access consciousness*.

Phenomenal consciousness may yet be due to activity in area V1.

The morale: common “folk-psychological” concepts may well turn out to be too crude in the light of scientific evidence.  $\implies$  *eliminativism* (the Churchlands). There is no such thing as “consciousness”.

Thoughts do not literally exist. Cf. human memory: psychologists make a distinction between several totally different types of memory.

Recent problem, studied by both neuroscientists and philosophers:  
“binding problem”. Is “binding” due to synchronized activity in the  
30–50 Hz range?

Other interesting issues:

- Interpretation of modern imaging techniques (PET, fMRI, MEG). Do they make thoughts visible?
- Blindsight (morale: as above), reverse blindsight.
- Split brains (two persons in one body?) (evidence against Cartesianism).
- Kornhuber's readiness potential and the freedom of the will (evidence against Cartesianism?).

conventional picture:

volition – brain activity – voluntary movement

experimental evidence (?):

brain activity – volition – voluntary movement

## **5 Are the neurosciences fundamentally deficient?**

Mary (Frank Jackson)

## Part VII

# The hard problem

The problem of consciousness.

My best guess: two perspective theory, internal perspective vs. external perspective (p. 55).

Is the hard problem *too* hard for us? (McGinn)

*Ignoramus, ignorabimus?*

These transparencies are available at

<http://www.eur.nl/fw/staff/lokhorst/helsinki.2001.html>