

## Chapter II

# Materialism vs. spiritualism

## Does a perfect human material copy have identical consciousness?

### 1. Human duplication thought experiment

Consider the following thought experiment, typical of science fiction, involving the creation of an exact material copy of a person.<sup>17</sup> The person to be copied will be called Calvin-1, and its copy, Calvin-2, is built by recreating almost instantaneously each molecule of Calvin-1, in such a way that the state of each molecular replica is the same as that of the original molecule in Calvin-1, and all of the spatial relations (and also other types) between the molecules of Calvin-2 are the same as the relations between the molecules of Calvin-1.



Figure II.1. Calvin & Hobbes's invention.

When the reproduction takes place, Calvin-1 and Calvin-2 are in perfectly similar environments, so that in the first moments their material states will be perfectly similar, at least down to the molecular level. After a few moments, however, the two systems would start to roam in different directions, because of the inevitable differences in the fluctuations of each environment, or because nature is not deterministic (or both).

This is thus the setup of the thought-experiment. The first question to be asked is if Calvin-2 has consciousness, or if he is only a “zombie”, acting on automatic reactions but without subjective experience. Materialists would say that he is conscious, since consciousness is considered to be produced only by matter, while spiritualists or “substance dualists” (such as Descartes) would claim that something more would be needed for Calvin-2 to have a soul, a mind, or consciousness.

### 2. Materialism and physicalism

*Materialism* is the view according to which what we call soul, spirit, mind or consciousness is only the product of material processes, and that upon the death of the body the individual mind in fact disappears. In its broader sense, materialism is not committed to the claim that “matter” is the fundamental substance of the universe, but it tends to be committed to the thesis that the fundamental entities of the universe (be them particles, fields, energy, strings, or whatever) are inanimate, without purposes or other mental attributes.

Materialism has a long history, starting with Greek-Roman atomism and the *carvaka* school in Ancient India. In the 17th century, Thomas Hobbes was a mechanistic materialist, and in the French Illustration one may mention Julien de la Métrie, with a more vitalistic

<sup>17</sup> This situation is presented in KIM, J. (1982), “Psychophysical supervenience”, *Philosophical Studies* 41: 51-70. It is practically the same as the “new scanner teletransporter” explored by Derek Parfit in his discussion on personal identity: PARFIT, D. (1984), *Reasons and persons*. Oxford: Oxford University Press, pp. 199-201. Concerning human material copies in science fiction, one may mention the film *The 6th Day*, from 2000, starring Arnold Schwarzenegger. The machine built by Calvin appears in the cover of: WATTERSON, B. (1991), *Scientific progress goes “boink”*, Andrews & McMeel, Kansas City. See also discussion in KURZWEIL (2005, op. cit., note 4, pp. 383-6).

materialism.<sup>18</sup> With the rise of the science of physiology, materialism gained new strength in the German speaking countries around 1850, being later surpassed by Kantism and positivism (in the philosophy of science), but generating dialectical materialism. Of the two major problems of materialism, one of them, the problem of the perfection of life, has found an adequate solution with the theory of biological evolution, but the problem of explaining how consciousness arises from matter remains unsolved. In the 1950's, authors such as Place, Feigl and Smart took up the materialist perspective in the philosophy of mind, exploring the mind-brain identity thesis, to be explored in Chap. IV.

The term *physicalism* is roughly synonymous to materialism. In the philosophy of mind, “physicalism” may be defined (in its ontological sense) as the claim that everything has a physical nature, including the mind. The problem with this definition is that it is based on the definition of “physical”, and there isn't a consensus on that. For now, we may characterize a physical process by means of three properties: (i) location in space and time; (ii) occurrence at a scale (micro, macro, etc.); and (iii) the absence of final causes, i.e., at the elementary level there are no purposes or intentionality.

### 3. Determinism and mechanicism

One says that at a certain moment a physical system is in a certain *state*. This term usually involves an ontological component, expressing a set of real properties of the system, but it may also express the epistemological limitations that the observer has in relation to the system, which is usually never fully known.

Over time, the state usually changes, and one can speak of the “evolution” of the system, in the sense that it is simply having its state change. A fundamental question regarding any system is whether the state, at a certain moment in time, fixes uniquely the future states. If so, the system is said to be *deterministic*.

Determinism is a concept that involves the notion of *causality*. If we say that “an event A causes an event B”, we can understand this as the statement that “in the absence of A, B would not occur”. If this is true, we would have that A is the *necessary* cause of B. On the other hand, it could be the case that in the absence of A, B would continue to occur, because another factor A' is present, which is also cause of B. In this case, if A (besides A') is sufficient for the occurrence of B, we say that A is a *sufficient* cause of B.

If a system is deterministic, one says that a present state is a sufficient cause for a future state to take place. If it is also a necessary cause, one says that the system is also “reversible”.

An important thesis that will appear in discussions about the relation between the physical and the mental is that the physical domain would have the property of “causal closure”, in the sense that every physical event has a set of sufficient causes that determine it. However, one must not impose that there is strict determinism in the physical world, that is, that the past state of the universe uniquely determines the future. The problem of whether

---

<sup>18</sup> On the history of materialism, see LANGE, F.A. ([1875] 1974), *The history of materialism*, 3 vols. in one, transl. E.C. Thomas (1879-81), Arno Press, New York (1<sup>st</sup> German: 1866). On the carvaka doctrine (pronunciation: cháravaka), see DASGUPTA, S. ([1922] 1975), “The lokayata, nastika and carvaka”, in *A history of Indian philosophy*, vol. III, Motilal Banarasiidars, Delhi, pp. 512-50. A comparison between the mechanistic materialism of Hobbes and the French vitalistic materialism (of Julien de la Métrie and Diderot) appears in SKRBINA, D. (2005), *Panpsychism in the West*, MIT Press, Cambridge, pp. 101-5. Concerning the German materialism of Vogt, Molleschott, Büchner e Czolbe, see also GREGORY, F. (1977), *Scientific materialism in nineteenth century Germany*, Reidel, Dordrecht. There is a lot about post-war materialism, for example: MOSER, P.K. & TROUT, P.K. (eds.) (1995), *Contemporary materialism: a reader*, Routledge, London.

nature is deterministic or “tychist” (indeterministic) remains an open question: even in quantum physics, where individual experimental results are unpredictable (what is predictable is the statistics of a large number of results), there are deterministic interpretations such as David Bohm’s (1952).

Another characteristic to be considered is whether the description of a physical domain is purely quantitative, given in mathematical and geometrical language, or whether it involves “qualities”. We will apply the term *mechanicism* (or “mechanism”) to the purely quantitative description, involving only “figure and motion” (Leibniz, *Monadology*, § 17). Classical mechanicism of the days of Descartes and Hobbes assumed that only collisions could make a body change from its inertial state. Newton introduced forces that act at a distance, while Leibniz and Boscovich proposed the notion of a “center of force” that would replace the idea that matter fills space uniformly. In the 19th century, the so-called fall of the mechanical worldview led to new principles not encompassed by Newtonian mechanics, such as the principle of relativity (for all physical processes) and the principles of quantum physics. Notwithstanding, physicists continue to speak of “quantum mechanics” and “relativistic mechanics”. So I will define “mechanicism” in its broadest sense, as the supposition that the theoretical description of physicists, by means of equations of motion, exhausts all there exists in physical reality.

An example of a non-mechanistic approach is to consider that the qualia (section I.6) that we experience subjectively are part of the physical processes, constituting real and qualitative aspects of the world. An extension of this “qualityism” to all of reality would result in a non-mechanistic physicalist worldview.

#### 4. Spiritualism and parapsychology

Most religions teach that upon death our individual soul survives in a supernatural realm, being or not able to transmigrate to other living beings in the future. This conception may be called *spiritualism*, and includes “moralist” positions, in which the soul is conceived as entering into a state of unconscious sleep (Luther) or death (John Milton), until the resurrection of the soul.

In section A1.1 we present Plato’s dualistic view. Interestingly, the thesis of the existence of incorporeal souls was supported by some empirical evidence, such as the example of the will-o’-the-wisp (*ignis fatuus*) in graveyards. Nowadays, there is a debate between materialists and spiritualists surrounding the truth of the parapsychological theses. An example is the *near-death experience*, in which people who were on the verge of death, but survived, report a similar set of experiences, such as a peaceful sense of having died, a review of their entire life history, the vision of a tunnel with light at its end, and the experience of leaving the body. Is the experience of near-death explained by neurological processes, or is it a clue to the existence of a supernatural world? Materialists and spiritualists are divided on the answer.

There are many experiments in which parapsychologists report favorable evidence for so-called “psi phenomena,” such as telepathy (transmission of thought) and premonition (predicting the future), but most orthodox scientists who embrace what might be called “scientific skepticism” consider them a fraud or the fruit of self-deception, which would occur especially in the phase of data collection. The discussions usually end by invoking experiments favorable to parapsychological theses, which skeptics do not accept, and which orthodox scientists have no patience to try to reproduce and falsify.

## 5. The analogy of clocks and mirrors

What is the relation between mind and body? In section A1.6, we present the dominant views of the 17th century, in Europe, which included Descartes' interactionism, Malebranche's and others' occasionalism, and Leibniz's pre-established harmony.

The Belgian Arnold Geulincx, in the "Annotations" (§ 19) to his work *Ethics*, published with the book in 1675, proposed an analogy between two clocks to illustrate interactionism and occasionalism (his personal view). Leibniz was indirectly influenced by Geulincx's text,<sup>19</sup> and developed the analogy in 1696:

Imagine two clocks or watches which are in perfect agreement. Now this can happen in three ways. The first is that of natural influence. This is the way with which Mr. Huygens experimented, with results that greatly surprised him. [...] The second way of making two clocks, even poor ones, agree always is to assign a skilled craftsman to them who adjusts them and constantly sets them in agreement. The third way is to construct these two timepieces at the beginning with such skill and accuracy that one can be assured of their subsequent agreement.

Now put the soul and the body in the place of these two timepieces. Then their agreement or sympathy will also come about in one of these three ways. The way of influence is that of the common philosophy [and of Descartes' interactionism]. [...] The way of assistance is that of the system of occasional causes [Malebranche's occasionalism] [...] Thus there remains only my hypothesis, that is, the way of pre-established harmony, according to which God has made each of the two substances from the beginning in such a way that though each follows only its own laws which it has received with its being, each agrees throughout with the other [...].

In 1860, the German philosopher Gustav FECHNER<sup>20</sup> generalized this analogy to include the aspect dualism of Spinoza:

Leibniz has left out one point of view – the most simple possible. They can keep time harmoniously – indeed never differ – because they are not really two different clocks [but only one]. We can therefore dispense with the mechanism of interaction, the constant adjustment, and the artificiality of the pre-established arrangement.

In this case, we can imagine a single clock and reflections of its image in two different mirrors (one of them would be the point of view that the clock has of itself). This is a didactical analogy in order to distinguish monist views (a single clock, as in the views of Spinoza and Fechner) and dualist ones (two clocks, as in Descartes, Malebranche and Leibniz).

With the use of mirrors, the clock analogy can be extended to other monistic conceptions. The defense of the mind-body identity thesis by Australian materialism, as we will see in U.T. Place, in fact privileges the material description, which in this new analogy would be represented by the clock, while the mind would be a reflection in the mirror (subject to "phenomenological fallacy"). This would express a form of reductionist materialism.

---

<sup>19</sup> See more details in the text available in the website of the course, on Leibniz's clock analogy. The quote is from LEIBNIZ, G.W. (1696), "Second Explanation of the New System", postscript of a letter written to Henri Basnage de Bauval, written in Jan. 3-13, in LOEMKER, L.E. (ed.), *Gottfried Wilhelm Leibniz: philosophical papers and letters*, 2nd ed., Reidel, Dordrecht, 1976, p. 459.

<sup>20</sup> FECHNER, G.T. (1866), *Elements of psychophysics*, vol. I, transl. H.E. Adler; Holt, Rinehart & Winston, New York, (German original: 1860), p. 4.

Epiphenomenalism, which we will see later, can also be fit into this model of a material clock (with causal powers) and a mental mirror (without its own causal powers).

Reversing this latter model would lead to a variety of idealism: the material world would be a mirror reflection of the mental clock.

The Aristotelian thesis that everything has form and matter could perhaps be interpreted by the model of a single clock with no mirror: it would not be two different perspectives of the same thing, but one thing containing two kinds of properties (“causes”): form and matter. The “colored brain thesis” could also be represented by a single clock, without mirrors: the clock would consist of qualia and other mental contents, as well as of the structural properties described by science.

## 6. Psychophysical parallelism

The term “psychophysical parallelism” or “psychophysiological parallelism” was very much used in the period between 1860 and 1930, to denote the correlation between mental events and bodily events. “The theory that the conscious and nervous processes vary concomitantly whether or not there be any causal connection between them”; “each psychical change or psychical state, each *psychosis*, involves a corresponding neural change or neural state, *neurosis*, and vice versa”.<sup>21</sup> This may be called PARALLELISM<sub>1</sub>.

A restricted meaning, PARALLELISM<sub>2</sub>, was used in Britain, adding the clause that no interaction took place between the mental and the bodily series, thus referring to views such as Leibniz’s pre-established harmony.

Analyzing the work of Fechner, the German historian Michael Heidelberger considers that, besides the “empirical postulate” of meaning 1, there is another sense of psychophysical parallelism that Fechner called “the identity view”: PARALLELISM<sub>3</sub>. This is Fechner’s “doctrine of the two perspectives”, close to Spinoza’s aspect dualism. For the pioneer of psychophysics, a human being is a single entity, whose properties are considered mental when perceived internally, and considered physical when viewed externally. Fechner introduced the expression “parallelism of the mental and the physical” in 1861.<sup>22</sup>

The expression “psychophysical parallelism” (sense 1) started being used to express the recognition that the science of physiology had in fact established many of the bodily concomitants of mental processes, and by induction this could be extended to *all* mental processes. This is clear in the following quotation by the Scottish philosopher and psychologist Alexander Bain:<sup>23</sup>

Thus we have physiological evidence on the one hand, that a certain time is occupied by the nerve-force, and we have mental evidence on the other, that an equivalent time is occupied by sensation, thought, and volition. Our [38] thinking can never transcend the physical pace of the nerve-force (pp. 37-38).

<sup>21</sup> The first quotation is from the entry “Parallelism, Psychophysical”, of the *Encyclopædia Britannica* of 1911, vol. 20, p. 762. The second one is from the Jesuit WALKER, L.J. (1913), “Psycho-physical parallelism”, *Catholic Encyclopedia*, vol. 11. Both are available in the internet.

<sup>22</sup> HEIDELBERGER, M. (2004), *Nature from within: Gustav Theodor Fechner and his psychophysical worldview*, trad. C. Klohr, U. Pittsburgh Press, ch. 5 (German original: 1993). The use of the term “parallelism” by Fechner, in 1861, is indicated in p. 101. This chapter is also available as a separate paper in the internet: HEIDELBERGER, M. (2001), “The mind-body problem in the origin of logical empiricism: Herbert Feigl and psychophysical parallelism”, *PhilSci Archive* 945, 26 pp.

<sup>23</sup> BAIN, A. (1873), *Mind and body: the theories of their relation*, H.S. King & Co., London.

The simplest term that we can employ for a mental state is a *shock*; a word equally applicable to the bodily side and to the mental side. [...] because there is a rapid transition from quiescence to excitement; in which circumstance there is an accurate parallelism between the otherwise distinct physical and mental facts. (p. 40).

Parallel to this mental series is the physical series of facts, the successive agitation of the physical organs [...] While we go the round of the mental circle of sensation, emotion, and thought, there is an unbroken physical circle of effects. [...] When, therefore, we speak of a mental cause, a mental agency, we have always a *two-sided cause*; the effect produced is not the effect of mind alone, but of mind in company with body. (p. 131).

Notice that the concept is *symmetrical* in relation to the perspectives of mind and body. The stipulation that there be a one-to-one correspondence between sensations and “psychophysical” states was explicitly formulated by George Elias Müller (1896), who pointed out as his precursors Ernst Mach (1866) and Ewald Hering (1878). Following this tradition, the Gestalt psychologist Wolfgang Köhler introduced the term “psychoneural isomorphism” in 1929,<sup>24</sup> to express the correlation between systemic (structural, functional) properties of mind and brain. Only Hering appears to have added to this a correspondence between subjective qualities and chemical qualities in the nervous cells (which he called “specific energies”).<sup>25</sup>

Psychophysical parallelism was criticized by many thinkers from the humanities, like Walker, quoted above:

The parallelist [...] asserts that intellectual operations have an exact physiological counterpart, which is more than he can prove. [...] But that intellectual operations proper – judgment, logical inference, general concepts, vast and far-reaching as they are in their significance, should have an exact counterpart in the activity of brain-cells and their neuron connections, is a hypothesis which the known facts of psycho-physics fail to bear out, and which is also inconceivable. How, for instance, can a general concept, referring as it does to objective reality and embracing schematically in a single act many diverse notes, bear any resemblance to the disturbance of nervous equilibrium that accompanies it, a disturbance which has no unity at all except that it occurs in different parts of the same brain more or less simultaneously?

This philosophical reaction against psychophysical parallelism began, in the German context, with the attacks by philosopher Christian Sigwart, followed by other critics, such as Wilhelm Dilthey, who pointed out that the position was too close to materialism, subjecting the mind to determinism, e taking away from the human soul the capacity to freely choose its causal actions upon matter. The defense of parallelism and of experimental psychology was made by the psychologist Hermann Ebbinghaus, with an ensuing debate with Dilthey (HEIDELBERGER, 2004, p. 179-80). The psychologist Wilhelm Wundt (1894) accepted parallelism in relation to elementary sensations and feelings, and their associations, but rejected it for higher mental functions (SCHEERER, 1994, p. 185). The chasm between philosophers and experimental psychologists would lead in the attacks against “psychologism” on the part of Frege, Husserl, and others.

<sup>24</sup> The term “psychoneural isomorphism” was introduced in English by KOHLER, W. (1929), *Gestalt psychology*, New York: Liveright. See historical and conceptual discussion in SCHEERER, E. (1994), “Psychoneural isomorphism: historical background and current relevance”, *Philosophical Psychology* 7: 183-210, and in BORING, op. cit. (note 8), pp. 83-90, 95-96.

<sup>25</sup> HERING, E. (1913), “On the theory of nerve-activity”, in *Memory: lectures on the specific energies of the nervous system*, 4<sup>a</sup> ed., Open Court, Chicago, pp. 43-70 (transl. of speech given in 1898 in Leipzig).

Walker's text also points out an inverse deficiency in the concept of parallelism, which is the fact that bodily changes frequently don't have an observable parallel in the mind. One way out for the parallelist was to invoke unconscious mental changes for every body process, even if microscopic, which would lead to conceptions in the panpsychist spectrum.

This problem would be assimilated by the *asymmetrical* concept of "supervenience", which will be now examined.

## 7. Supervenience of the mental upon the body

Let us return to the human duplication thought-experiment that opens this chapter. We assumed that a perfect human material copy of a person was made, call him Calvin-1, and we asked whether his copy, Calvin-2, would be conscious. Let us now adopt the materialist answer and examine its consequences; that is, let us assume that Calvin-2 is created with consciousness. The spiritualists who might want to follow the reasoning may consider a modification of the thought-experiment, and suppose that a soul or spirit was created together with Calvin-2.

Assuming that the material states of the two persons are perfectly similar down to the molecular scale, the second question to be asked concerns the nature of the mental states of Calvin-1 and Calvin-2 at the moment of reproduction. In this moment of creation, would the two be in identical states of consciousness? Would they have perfectly similar thoughts, the same dreams, emotions, would they be aware of exactly the same colors, and pay attention to the same smells?

The position that both consciousnesses would be perfectly similar is shared by most materialists, and expresses the *supervenience* thesis of the mental states upon the bodily states. The disagreement within the materialist camp is whether the mind "reduces" to the body, or whether it "emerges" as something partially independent (this distinction will be discussed in Chap. VII).

The supervenience thesis of the mind upon the body claims that the physical state of the body (the "subvenient level") fixates uniquely the mental state (the "supervenient level"). In other words, any change at the higher level, involving mental states, requires a change at the lower level, a physical or material change. Or, alternatively, if a lower level state does not change, then the higher level will also remain the same. This last condition is exactly what is explored in the human duplication thought experiment. The thesis that Calvin-1 and Calvin-2 have perfectly similar subjective experiences or mental properties is a consequence of the supervenience thesis of mental states upon material states of the human body.

Fig. II.2 illustrates the many-to-one mapping of brain physical states onto mental states, required by the supervenience thesis. If the physical states of the brain be considered at a molecular resolution, then of course small variations in physical states will not lead to a subjectively noticeable conscious modification. However, if a coarse-grain resolution is considered which groups all the physical states subvenient to a mental state (for example, states  $p_1$  to  $p_5$  in the figure) into a single physical state  $p$ , the relation between physical and mental states will approximate a one-to-one mapping. In what scale would this happen? This is an important empirical question which is still open. At such a scale, one can speak appropriately of psychophysical "parallelism" or "isomorphism".

In order to explore the supervenience thesis, consider the Neckar cube of Fig. II.3. It is a two-dimensional figure, but our mind is able to represent to itself, to "project", a figure which is similar to what we see when looking at a three dimensional cube. However, this projection can take place in two different ways, depending on which square "pops out of the paper". The transition between the two cases can be caused by conscious decision, but it can

also happen spontaneously. The point of this example is that a change in perception implies a change in the mental state. What consequence can be derived from this situation, using the supervenience thesis?

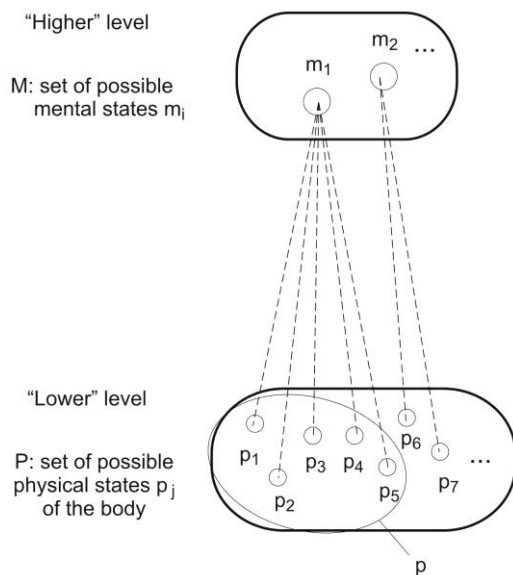


Figure II.2. Sketch of the supervenience relation of mental states  $m_i$  upon physical states  $p_j$ .

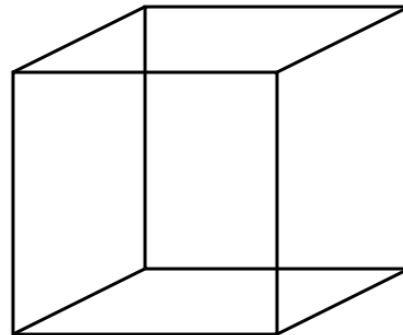


Figure II.3. Necker cube. From this two-dimensional figure (a non-simple dodecagon, which results from the oblique parallel projection of a cube), one can “mentally project” two distinct representations, according to which side of the cube appears to pop out of the paper.

The consequence is that necessarily there has to be a change in the *physical* state of the brain, when passing from one perceptual state to the other. When we perceive the left square popping out, and consciously change our perception so as to see the right square popping out, we necessarily change the physical state of the brain (according to the supervenience thesis).<sup>26</sup> In this sense, the mind has causal powers over matter! (But we still have to study whether the mind is itself a product of matter.)

## 8. The supervenientist spectrum

The supervenience thesis may be interpreted in different ways, according to the choice of the “minimal basis of supervenience”. A cerebralist says the mind supervenes upon the cerebrum or the brain. With this she means that what she is experiencing in her consciousness is not being produced immediately in the outside world, but immediately in her brain. We can only perceive a sailboat in the bay when the sunrays reflect on the boat and fall on our retinas,

<sup>26</sup> Compare with the discussion in WEYL, H. ([1927] 1949), *Philosophy of mathematics and natural science*, transl. O. Helmer, Princeton U. Press, p. 26: “It would be folly to expect cognition to reveal to intuition some secret essence of things hidden behind what is manifestly given by intuition. The idea of isomorphism demarcates the self-evident insurmountable boundary of cognition. This reflection has enlightening value, too, for the metaphysical speculations about a world of things in themselves behind the phenomena. For it is clear that under such a hypothesis the absolute world must be isomorphic to the phenomenal one (where, however, the correlation needs to be unique only in the direction thing in itself  $\rightarrow$  phenomenon); for ‘we are justified, when different perceptions offer themselves to us, to infer that the underlying real conditions are different’ (Helmholtz, *Wissenschaftliche Abhandlungen*, II, p. 656). Thus even if we do not *know* [*kennen*] the things in themselves, still we have just as much *cognition* [*Erkenntnis*] about them as we do about the phenomena.”



generating a chain of neurophysiological impulses which only becomes conscious after entering the physical brain, or a specific region of it, as claimed by localizationists (in relation to consciousness). So a cerebralist accepts that a cultural event can shape our consciousness, having a causal effect upon our mind; but this can only happen after sensorial information enters the brain.

Many, however, argue that this physicalist approach that tracks down the transmission of information at every millisecond is excessively analytic, and that it is not possible to divide a person into small spatial and temporal parcels. These thinkers therefore assert that consciousness supervenes minimally in more extended regions than the brain. The approaches included in the “4E cognition” (*embodied, embedded, extended, enactive cognition*) claim that the basis of supervenience should include the whole body, or even the environment around us that we use in our thinking. In the social sciences the notion that an individual does not exist without a society or culture is quite strong, so this can be used in favor of the claim that individual consciousness supervenes upon society or culture.

Table II.1 summarizes different positions in what may be called the “supervenientist spectrum”. It includes two negations of the thesis that the mind supervenes upon some physical system. A spiritualist like Henri Bergson denies that the same physical state generates a single mental state because he conceives of the spirit as having a much broader non-material dimension (last line of table). On the other hand, a materialist might deny that Calvin-1 and Calvin-2 are (at the moment of creation) in the same mental state because they simply deny that the relation of determination from the physical to the mental (or from micro to macro) is strict (first row of table).<sup>27</sup>

CONSCIOUSNESS SUPERVENES MINIMALLY TO WHAT?	CONCEPTION
To nothing	Anti-supervenientism
To a specific region of the brain	Localizationism of consciousness
To the whole brain	Brain holism (cerebralism)
To the body of the animal	Embodied cognition
To the body and the environment that we manipulate	Extended cognition
To all of society and culutre around us	Culturalism
To the whole material Universe	Universal holism
To an extra-material entity (even if partially)	Spiritualism

Table II.1: The supervenientist spectrum. Different materialist positions, concerning the issue of what produces consciousness immediately (only the last line denies materialism).

### 9. Internalism vs. externalism

The Korean-American philosopher Jaegwon Kim (1982)<sup>28</sup> explored many examples of *relational properties* that distinguish Calvin-1 from Calvin-2, at the moment of creation. To

<sup>27</sup> In the case of causality, something similar happens for worlds in which strict determinism is not valid. Even though it is counterintuitive to imagine that an event can arise spontaneously, without being uniquely determined by a set of causes, this is a situation considered possible by contemporary physics. Similarly, even though it is counterintuitive that the nanoscopic scale does not uniquely fix the macroscopic, this may indeed happen, according to the anti-supervenientist.

<sup>28</sup> KIM (1982), op. cit. (note 14).

begin with, they are in different positions in space: however, this does not seem to be important for characterizing their mental states.

But suppose that Calvin-1 was in fact in a party the night before the experiment, a party in which Calvin-2 could not have been. After the creation of Calvin-2, both could be in perfectly similar mental states, thinking about the party with equal richness of details. (An alternative analysis could argue that, given that Calvin-2 was generated by a causal process from Calvin-1, then he would share the identity of Calvin-1 before the experiment, so that one could say that he also was at the party on the previous night.)

Kim (1982, pp. 57-8) makes a list of some psychological properties that would not supervene on physical properties. We will mention some of them, adapted for Calvin.

(1) Calvin-2 thinks he remembers the party, but in fact he doesn't. Or better, Calvin-1 "knows" that he went to the party, and "truthfully believes" in this. Calvin-2 believes falsely.

(2) Calvin-1 is now happy *for having been* invited for the party. Calvin-2 is now also happy, but not "for having been invited", because he was not invited (only Calvin-1 was).

(3) At the moment of being copied, Calvin-1 saw a photograph of Hobbes in his chamber. The chamber of Calvin-2 was built in a perfectly similar way, but the photograph in his chamber is numerically distinct from the first. Therefore, they *are not seeing* the same object, even though their perceptual mental states are perfectly similar, and they are in "the same appropriate neural state" (p. 58).

(4) If in the moment of creation Calvin-1 signed a check to pay for the present he took to the party, Calvin-2 would be signing a check, but not to pay *his* present. Strictly speaking, he wouldn't even be "signing a check", since he is not yet a member of society, doesn't have a bank account, etc.

Such examples can be used as criticism to a "cerebralist" position, and in favor of a "culturalist" one, that there is only supervenience of mind upon the whole of society, past and present. Such positions are examples of an important division in the philosophy of mind, between *internalist* and *externalist* points of view.

For a *mental properties* (or mental content) *internalist* (the cerebralist or versions of embodied cognition), the truth value of a proposition (taking case 1 above) is not part of the relevant attributes of a mental state. The truth value is a *relation* between a belief and the state of affairs of the outside world, and not an intrinsic attribute of a belief or of a proposition held in the mind.

On the other hand, for an *externalist* (the "culturalist" or "holist" of Table II.1), a change in the world can lead immediately to a change in the mental state (without there being propagation of physical information between them), i.e., the mental state depends immediately on parts of the natural and social that surrounds the person.<sup>29</sup>

This distinction illuminates an anecdote regarding a case that occurred at the II Brazilian Meeting on Brain and Cognition (UFABC) in September 2013, in which the philosopher João Teixeira provided an argument against reductionist materialism, according to which a proposition in our mind is either true or false, but a neuron or part of the brain tissue is neither true nor false. This generated an inflamed reaction in a foreign neuroscientist! The difference between both thinkers boils down to the distinction between an externalist view and an internalist view of the mind.

---

<sup>29</sup> The classic defense of mental content externalism is PUTNAM, H. (1975), "The meaning of 'meaning'", in Gunderson, K. (org.), *Language, mind and knowledge*, University of Minnesota Press, Minneapolis, pp. 131-93.