Abstract Eliminative materialists argue that we can overcome the phenomenological gap between two different ways of referring to our subjective experiences – either as introspectively grasped in terms of folk psychology or as explained in neurological terms – by abandoning the pre-scientific concepts of folk psychology. However, unless these theorists can offer a plausible explanation for why the scientific view of the human mind proposed by cognitive neuroscience is so deeply counter-intuitive, this argument will remain unconvincing. In order to address the difficulties involved in making the paradigm shift from folk psychology to cognitive neuroscience I (a) briefly review the theoretical revolution that marked the transition from classical mechanics to the theory of relativity at the beginning of 20th century; (b) identify some similarities between this paradigm shift in physics and the birth of a new scientific view of the mind; (c) explain by means of (a) and (b) why neurological theories that reduce consciousness and the Self to aspects of brain dynamics appear implausible from a common sense perspective despite being sound from a scientific point of view.

KEYWORDS: Time; Consciousness; Relativity Theory; Neuroscience; Eliminative Materialism.

Riassunto Tempo e coscienza nel naturalismo cognitivo - I materialisti eliminativi sostengono che il divario fenomenologico tra due modi diversi di riferirsi alle esperienze soggettive, ora colte introspettivamente nei termini della psicologia del senso comune e ora spiegate in termini neurologici, può essere superato abbandonando i concetti pre-scientifici della psicologia del senso comune. Tuttavia essi continueranno a essere poco convincenti fino a quando non riusciranno a formulare una spiegazione plausibile del perché l’immagine scientifica della mente umana offerta dalla neuroscienza cognitiva sia così profondamente contro-intuitiva. Per superare le difficoltà legate al cambiamento di paradigma dalla psicologia del senso comune alla neuroscienza cognitiva si presenta qui un argomento articolato in tre passaggi: (a) viene brevemente esaminata la rivoluzione teorica che ha segnato il passaggio dalla meccanica classica alla teoria della relatività all’inizio del XX secolo; (b) vengono messe in luce alcune somiglianze tra, da un lato, il summenzionato cambiamento di paradigma in fisica e, dall’altro, la nascita odierna di una nuova visione scientifica della mente; (c) viene spiegato per mezzo di (a) e (b) il motivo per cui le teorie neurologiche che riducono la coscienza e l’io ad aspetti della dinamica del cervello appaiono implausibili al senso comune, sebbene siano pienamente convincenti da un punto di vista scientifico.

PAROLE CHIAVE: Tempo; Coscienza; Teoria della Relatività; Neuroscienze; Materialismo Eliminativo.

S. Nannini - Dipartimento di Scienze Sociali, Politiche e Cognitive, Università di Siena, Palazzo San Niccolò, via Roma, 56 - 53100 Siena (I)

E-mail: sandro.nannini@unisi.it

Creative Commons - Attribuzione - 4.0 Internazionale
The Hard Problem

Since the birth of the cognitive sciences fifty years ago, the boundary between the natural sciences on the one hand and philosophical reflections on the nature of human beings and their world on the other hand has become increasingly permeable.

From antiquity until very recently questions like What is the mind? What is consciousness? What is the Self? What is the relation of the mind to the body? Do human beings have free will? have been the exclusive preserve of philosophers and theologians. But nowadays, such questions are more and more frequently debated by experimental psychologists, computer scientists and neuroscientists.

However, although this “cognitive turn” is a new cultural trend of our age, the way in which this has been interpreted by philosophers of mind echoes to a large extent three solutions offered for the mind-body problem from antiquity to the present day, that is, dualism, materialism, and functionalism. Criticisms made by followers of each of these proposed solutions of the others also echo earlier rebuttals. For example, according to dualists, materialists incur a category mistake when they identify phenomenal consciousness with the brain activity from which it emerges.

Explaining how phenomenal consciousness emerges from brain activity is seen by many philosophers of mind and cognitive scientists as the Hard Problem par excellence. Many of them believe indeed that this problem is not only hard but also insoluble in principle. Materialists – and particularly the most uncompromising ones among them, often called “eliminativists” – argue on the contrary that the phenomenological gap between two different ways of referring to our subjective experiences – as introspectively grasped in terms of folk psychology or as explained in neurological terms – can be overcome by abandoning the pre-scientific concepts of folk psychology by means of which mental states have hitherto been described and replacing them with new concepts taken from cognitive neuroscience: a new discipline that nowadays, according to eliminativists, is causing a real paradigm shift in the science of mind and consequently in the view that we human beings have of ourselves.

However, despite the great progress made by cognitive neuroscience over the last twenty years, the eliminativist argument will remain unconvincing until eliminativists are able to offer a plausible explanation for why the scientific view of human mind offered by cognitive neuroscience is so deeply counter-intuitive and appears to most philosophers and scientists to be completely incomprehensible. How to escape from this predicament? The solution I propose is the following one:

a) I will briefly review an exemplary case of paradigm shift: more exactly the theoretical revolution that marked the transition from classical mechanics to the theory of relativity at the beginning of 20th century.

b) I will try to identify some similarities between, on the one hand, this paradigm shift in physics and, on the other hand, the cultural process that for the last fifty years has been leading many philosophers and scientists all over the world to abandon a vision of human beings and their world still based on Descartes’ mind-body dualism for a new scientific conception of the mind as brain activity in interaction with the external world.

c) I will try to explain through (a) and (b) why the (allegedly platonic) image of the soul as a “sailor” able to drive the body appears to common sense introspectively more appropriate (and more useful in everyday life) than a neurological theory of consciousness and the Self that reduces them to aspects of brain dynamics, although, on the other hand, from a scientific point of view, the image of the Self as a self-conscious free entity ontologically independent of the body is not plausible.

If this research strategy, here only broadly sketched, succeeds then we might hope to solve the Hard Problem or at least to learn to
live more peacefully with it.

Time, space and gravity in the theory of relativity

The earthquake that Einstein caused in the physics of his time had its epicentre in a deep change in the concept of time. This change brought about other changes in all fundamental concepts of physics as a ripple effect. Thanks to this theoretical revolution Einstein managed to solve the main problem that haunted physicists at the end of 19th century: how can classical mechanics be reconciled with the theory of electromagnetism?

In order to clarify both this problem and Einstein’s solution in a very simple way that is understandable first of all to me (I am no physicist) let me introduce a science-fiction thought experiment.  

A spaceship travels at the constant speed \( v=0.8c \) from Earth to a space station located at the extreme periphery of the solar system (at a distance from Earth that always remains the same) and as soon as the spaceship reaches the space station it comes back to Earth at the same speed.

Two observers – \( O \) from the control tower of the space port on Earth and \( O' \) from within the spaceship – simultaneously launch a radio signal to the space station when the spaceship is taking off. \( O \) receives eight hours (8h) later the confirmation from the space station that her signal has reached its destination. Assuming that the signal travelled at the speed \( c=1 \) (that is, the speed of light and all electromagnetic waves in a vacuum if the space and time units of measurement are suitably chosen) both on the outward journey and on the return journey, \( O \) (a staunch supporter of classical mechanics) calculates that the space station is at a distance of four light hours (8h:2=4lh) from Earth and that the spaceship, travelling at 0.8c, should have reached the space station five hours (5h) after its departure from Earth (4lh:0.8c=5h).

Thus, by assuming that the spaceship maintains the same speed on its return journey (but in the opposite direction), \( O \) forecasts that the spaceship will be back on Earth ten hours (10h) after its departure. And in fact she can verify later that this is so according to her clock. In addition, by assuming the existence of absolute time in accordance with classical mechanics, \( O \) calculates that also according to the clock of \( O' \), the signal that \( O' \) sent at the time of the departure of her spaceship from Earth should have taken four hours to reach at speed \( c \) the space station and should have returned back to the spaceship taking about twenty-seven minutes (0.4h) on its return journey. Therefore, \( O' \) should have picked up the return signal from the space station about four hours and twenty-seven minutes after her departure from Earth (4h+0.4h= 4.4h).

Figure 1 summarizes the set of observations and predictions of \( O \) (according to classical mechanics). In this figure, \( K \) is the reference system of \( O \) (its \( x \)-axis represents the direction in which the spaceship moves to reach the space station whereas its \( t \)-axis represents the flow of time for \( O \)). \( K' \) is the reference system of \( O' \) as it appears in \( K \) (according to classical mechanics) if you assume that \( O' \) moves with respect to \( O \) by uniform rectilinear motion along the \( x \)-axis at the speed \( v=0.8c \). Therefore, \( K' \) is by definition an inertial system with respect to \( K \).

In other words, if time is absolute as is stated by classical mechanics then \( t^* \), the time-axis of \( K' \), must always remain parallel to \( t \), the time-axis of \( O \), regardless of the speed at which \( O' \) moves with respect to \( O \). As you can see in Figure 1, \( O' \) drags \( t^* \) behind herself, as it were, while she progressively goes away from \( O \). And this happens in such a way that \( t^* \) remains always parallel to \( t \).

It is only thanks to this parallelism of \( t^* \) to \( t \) that in accordance with classical mechanics every possible event has a time coordinate identical in \( K \) and in \( K' \). But in this way the signals emitted at the time of the spaceship’s departure from Earth no longer reach the space station with the same speed both for \( O \) and \( O' \).
Time and Consciousness in Cognitive Naturalism

Figure 1. K and K* are respectively the reference systems of the observers O and O'. The x-axis of K is coincident with the x*-axis of K* (the x*-axis, represented by an horizontal dotted line, is drawn above the x-axis only to distinguish them graphically). They represent the route followed by the spaceship during its travel from Earth to the space station (and return); x is a variable that measures the distance of any event (that is any point in the space-time plane identified by K and K*) from O (that is, from Earth) in K; x* measures the distance of any event from O' (that is, from the spaceship) in K*. The t-axis of K and the t*-axis of K* represent the passage of time respectively for O and O'. O represents also the origin of K (that is, the intersection point of its axes), O' represents also the origin of K*; t and t* are variables that measure, respectively in K and K*, the time elapsed between any event and the departure of the spaceship from Earth. The origins of K and K* are coincident in the moment in which the spaceship is taking off from Earth but the origin of K* slides along the x-axis insofar as the spaceship is getting away from Earth with velocity v. The event E1 is the arrival at the space station of the radio signals emitted by O and O’ in the moment in which the spaceship is taking off from Earth. E2 is the arrival of the spaceship to the space station. E3 is for O the position (in the space-time) of the spaceship in the moment in which the radio signals reach the space station. The amplitude of the angle α in K and the amplitude of the angle β in K* are proportional to the velocity of radio signals (that is, of light) respectively for O and O’. Therefore the “line of universe” along which radio signals travel from Earth to the space station is O-E1 for O in K and O’-E1 for O’ in K*. O-E2 is the “space-time trajectory” of the spaceship from Earth to the space station for O in K.

Figure 2. K and K’ are respectively the reference systems of the observers O and O’. The axes of K’ (that substitutes K*) are represented by dotted lines. Like in Figure 1 x is a variable that measures the distance of any event from O in K, x’ measures the distance of any event from O’ in K’. Similarly the (ic)t-axis of K and the (ic)t’-axis of K’ represent the passage of time respectively for O and O’ whereas t and t’ are variables that measure, respectively for O in K and for O’ in K’, the time elapsed between any event and the departure of the spaceship from Earth. E1 and E2 have the same meaning as in Figure 1 (E3 has been neglected). However, unlike in Figure 1 the origins of K and K’ are always coincident and the velocity v with which O’ is getting away from O is represented in K by the amplitude of the angle θ with which the (ic)t’-axis is clockwise rotated with respect to the (ic)t-axis. Also the x’-axis is rotated with the same angle θ with respect to the x-axis but this time anticlockwise. The velocity of radio signals (that is, of light) is identical in K and K’. Therefore, the “line of universe” followed by the radio signals travelling from Earth to the space station is just the same in K and K’: O-E1 is coincident with O’-E1. Also the “space-time trajectory” O’-E2 (or O’-E2) of the spaceship from Earth to the space station is identical in K and K’ but in K’ it lies on the “line of universe” of O’, that is, on (ic)t’.
This is very clear in Figure 1: the speed of the signals from Earth to the space station is measured by the amplitude of the angle $\alpha$ in $K$ and by the amplitude of the angle $\beta$ in $K^*$. The first angle is much larger than the second angle. Consequently, the speed of the signals is $c=1$ for $O$ both on their outward journey and on their return journey whereas their speed is for $O'$ $c-v=0.2c$ on the outward journey and $c+v=1.8c$ on the return journey.

The constancy of light speed, a central point of the theory of electromagnetism, is not compatible with the principle of classical mechanics that time is absolute, that is, it runs at the same pace in all possible reference systems (and therefore two events that are simultaneous for $O$ are simultaneous for $O'$ as well: see e.g. E1 and E3 in Figure 1).

However, $O'$ (who is a supporter of the Special Theory of Relativity (STR) and describes by means of Figure 2 how the whole situation appears to her) substitutes $K'$ for $K^*$ as her own reference system. In $K'$ the outward and return journey of the spaceship last only six hours ($6h'$) and not ten hours ($10h$) as in $K$. Moreover, $O'$ claims to have received the return signal from the space station about two hours and forty minutes ($2.5h'$) after her departure from Earth and not four hours and twenty-seven minutes ($4.3h$) as calculated by $O$. How is this possible?

The only way to explain the difference between the time measures$^8$ taken respectively by $O$ and $O'$ is to admit that one hour ($1h'$) of $O'$ corresponds to one hour and forty minutes ($1.5h$) of $O$: the time of $O'$ flows more slowly in comparison to the time of $O$. But this seems to be impossible!

According to classical mechanics assuming a slowing of time is absurd. Time is absolute. The constancy of light speed maintained by the theory of electromagnetism and confirmed by many experiments seems to be incompatible with classical mechanics. However, Einstein proved that such an incompatibility does not exist if the principle of the absoluteness of time is abandoned.

Figure 2 shows the solution recommended by STR to make classical mechanics compatible with the theory of electromagnetism. The two theories become compatible if you give up the absoluteness of time. This rejection of the absoluteness of time is represented in Figure 2 by replacing $K^*$ with $K'$. By this substitution the space-time trajectories (that is, the “world lines”) of the signals emitted by $O$ and $O'$ at the time of the departure of the spaceship from Earth become coincident (see the straight line on which the segment $O-E1$ lies). But to make this coincidence geometrically possible it is necessary to give up the parallelism between the $t$-axis (or better the $(ic)t$-axis) of $K$ and the $t'$-axis (or better the $(ic)t'$-axis) of $K'$.

The displacement of $O'$ with respect to $O$ is represented no longer by a shift that, with the passage of time, progressively moves the origin of $K^*$ at different points in $K$ along its $x$-axis (see Figure 1) but by a rotation of $(ic)t'$ of an angle $\hat{\theta}$ with respect to $(ic)t$ (see Figure 2). Consequently, since the constancy of light speed is represented in Figure 2 by the coincidence of the space-time trajectories of the signals respectively launched by $O$ and $O'$ (that is, by the “world line” of light on which the segment $O-E1$ lies) it is necessary that the straight line which represents this common trajectory not only be the bisector of the right angle formed by the $x$-axis with the $(ic)t$-axis in $K$ but also be the bisector of the angle formed by the $x'$-axis with the $(ic)t'$-axis in $K'$. Only thanks to the coincidence of these two bisectors, every event which lies in $O-E1$ has space and time coordinates (if the units of measurement are appropriately chosen) whose ratio is equal to 1 (that is, the light speed is $c=1$) in both reference systems.

However, this is possible only if, by renouncing the orthogonality of $K'$, after having clockwise rotated the $(ic)t'$-axis of an angle $\hat{\theta}$ with respect to the $(ic)t$-axis, you also rotate the $x'$-axis of an identical angle $\hat{\theta}$ with respect to the $x$-axis but this time anticlockwise. To sum up, representing the displacement of $O'$ with respect to $O$ by a double rotation (and not by a shift) of $K'$ with respect to $K$ is direct-
ly imposed by the assumption that the speed of light is constant for all observers in accordance with the theory of electromagnetism.

The characteristics that differentiate \( K' \) (in Figure 2) from \( K^* \) (in Figure 1) finally allow us to explain why according to STR it is impossible that the units of measurement of time intervals and space distances remain unchanged shifting from \( K \) to \( K' \). To find out how the units of measurement of spatial distances and time intervals vary in \( K' \) with respect to \( K, O' \) can rely on the coordinates of two privileged events in Figure 2: \( E1 \) (that is, the simultaneous arrival of the signals at the space station) and \( E2 \) (that is, the arrival of the spaceship at the space station). Since according to Figure 2 (that is, according to STR) both in \( K \) and \( K' \) the speed of the signals is \( c=1 \) whereas \( E1 \) lies on the world line of these signals, in both reference systems \( E1 \) has such coordinates that the difference between its space coordinate and its time coordinate is zero.

If the coordinates of \( E1 \) are \((x1\ t1)\) in \( K \) and \((x1'\ t1')\) in \( K' \) then the following formulas are valid: \( x1-t1=x1'-t1'=0 \) or according to a more general formulation\(^{10}\) (if you call \( s1 \) the space-time distance that separates \( E1 \) from \( O, \) that is, from the common origin of \( K \) and \( K' \)) the following formula is valid:

\[
(a) \ (s1)^2=(x1)^2-(t1)^2=(x1')^2-(t1')^2=0
\]

Moreover, \( O' \) knows that in \( K \) \( E2 \) is an event identified by the coordinates \( x2=4 \) and \( t2=5 \). In \( K' \) the space coordinate of the same event is instead \( x'=0 \). What is its time coordinate in \( K' \)? It may seem obvious by applying the Pythagorean Theorem that\(^{11}\)

\[
(b*) \ (s2)^2=(x2)^2+(t2)^2=(x2')^2+(t2')^2
\]

Because of the opposite signs of the squares of time coordinates “\(-\ (t1)^2\)” and “\(+\ (t2)^2\)” respectively in (a) and \((b*)\) these two formulas cannot be two different solutions of a more general single equation that determines the space-time distance \( O-E1 \) if the coordinates of \( E1 \) are \((x\ t)\) in \( K \) and \((x'\ t')\) in \( K' \). Since (a) is a consequence of the constancy of the speed of light whereas \((b*)\) is in accordance with Euclidean geometry and classical mechanics finding a way to make the two formulas become compatible is equivalent to making the theory of electromagnetism compatible with classical mechanics.

In STR this is possible by resorting to a mathematical trick devised by H. Minkowski, that is, by expressing time coordinates by means of imaginary numbers whereas space coordinates continue to be expressed by means of real numbers. Thanks to this trick \((b*)\) can be replaced (and generalised) by the following equation:

\[
(b1) \ s^2=x^2+(it)^2=(x')^2+(it')^2
\]

Thus, since – as is well known – the square of an imaginary number is a negative real number, from (b1) it is easy to obtain the following equation:

\[
(b2) \ s^2=x^2-t^2=(x')^2-(t')^2
\]

Unlike \((b*)\), \((b2)\) admits (a) as the special case in which \( s=0 \). Moreover, by reversing all signs except the sign of \( s^2 \) \((b2)\) can be transformed into the following equation:\(^{12}\)

\[
(c) \ s^2=t^2-x^2=(t')^2-(x')^2
\]

To understand why the travel of the spaceship lasted five hours \((5h)\) in \( K \) and only three hours \((3h^*)\) in \( K' \) is finally very easy by means of \((c)\) since by substituting in \((c)\) \( t, x \) and \( x' \) with the empirical data \( t=5, x=4, \) and \( x'=0 \) you obtain \((t')=3\):

\[
(d) \ \ (t')=(5^2-4^2)\text{/}2=3
\]

Moreover, if you generalize \((c)\) to all possible events by a simple algebraic calculation omitted here for brevity you obtain Lorentz Transformations.\(^{13}\) And by these transformations you can calculate all other values of...
Figure 2 (including how much slower time flows in \( K' \) with respect to \( K ((1h'=5h/3)) \). \(^{14}\)

To sum up, unlike classical mechanics STR can explain why \( O \) and \( O' \) take different measures of the same time intervals because it assumes that time flows for \( O' \) in a “direction” that is inclined by an angle \( \theta \) with respect to the direction in which it flows for \( O \). And the higher the speed with which \( O' \) moves with respect to \( O \) (or vice versa), the bigger that inclination. But to say that time flows for \( O' \) in a direction that is “inclined” with respect to the direction in which it flows for \( O \) is mysterious and incomprehensible to common sense. Time seems to flow only in one “direction” for all observers: from the past to the future through the present.

However, to make it less difficult to understand this crucial point of STR you can reflect on the fact that thanks to the constancy of the speed of light in all directions for all observers the distance covered by a ray of light in the time interval \( t \) is proportional to the space distance \( ct \) (space = speed \( \times \) time) in all reference systems. Therefore, you can replace – as is done in STR – all the measures \( t \) (or it if you use imaginary numbers) of time intervals in all reference systems with the corresponding measures \( ct \) (or \( ict \)) of the space distances travelled by light in those time intervals.

In other words, you can substitute the length of a time interval (in units of measurement that will change from one reference system to another) with the length of the distance travelled by light in that time interval. Of course, if the units of measurement of space distances and time intervals are chosen in all reference systems in such a way that \( c=1 \) it is irrelevant to write \( ct \) instead of \( t \) from the point of view of the results of mere calculations. \(^{15}\) But from a conceptual point of view that substitution is instead very important.

You can easily understand why it is so if you take into account that already in the first formulation of STR Einstein himself defined time by means of the procedure adopted to measure it. \(^{16}\) Therefore, if you measure and geometrically represent time intervals by means of the distances covered by light in those intervals you can, as it were, transfer the geometrical properties of the vectors that represent space distances to the vectors that represent time intervals. Consequently, the strangeness of a time that flows for \( O \) and \( O' \) in different directions (and therefore flows at a different rate) disappears if time is conceived as a fourth “space dimension” of a 4D-hyperspace (called the “space-time of Minkowski” in STR). \(^{17}\)

In other words, the core of Einstein’s theoretical revolution (if it is expressed through Minkowski’s geometrical representation) is that a vector representation of time intervals allows such a geometrical construction that the flowing of time for different observers in distinct directions (and therefore at a different pace) ceases to be absurd and incomprehensible.

Einstein himself developed Minkowski’s method to geometrically represent the properties of space and time as he formulated the General Theory of Relativity (GTR) by extending STR from inertial reference systems to accelerated systems. The tensor calculus required by GTR is too complex to be described here. Suffice it to say that GTR is a theory based on the so-called Equivalence Principle: the accelerated movements of a body in a reference system \( K \) become inertial movements in a reference system \( K' \) if this is appropriately accelerated with respect to \( K \).

Consequently, a body that moves with accelerated motion in \( K \) will move with inertial motion (or will stop) in \( K' \) if \( K' \) moves with accelerated motion with respect to \( K \) in an appropriate manner. For example, the objects inside the space station in orbit around Earth float (i.e. are stationary) for those who observe them from inside the station but they would appear to be in free fall toward Earth together with the whole space station for an observer on Earth who could observe them by telescope.

Moreover, at this point we have to take into account the fact that because of the time
dilation and space contraction imposed by STR the four axes of a reference system $K'$ that is accelerated with respect to $K$ are represented in $K$ no longer by straight lines but by Gaussian curves. Consequently, Lorentz Transformations (that are applicable only to inertial reference systems) are replaced in GTR by ten differential equations thanks to which the objective invariant physical properties of the gravitational field at a given point of space-time are expressed by the properties of its geometrical representation at that point (according to a geometry that is not Euclidean any more). The intensity of gravity at a certain point of space-time is not any more a force as in classical mechanics but rather a specific degree of space-time curvature at that point.

For example, in classical mechanics the motion of the Earth around the Sun is an accelerated motion in a “flat” Euclidean space indifferent to what happens in it. The acceleration is brought about by the attraction that the mass of the Sun mysteriously applies to the Earth and vice versa. In GTR the motion of Earth around the Sun is instead an inertial motion along a geodesic, that is, along the shortest path between two points lying on a non Euclidean space-time that is curved by the presence of masses in it.

To sum up, all fundamental concepts of classical mechanics – in particular time, space and gravity – are substituted in STR and GTR by other concepts that instead express physical properties by means of the geometrical properties of a vector representation of space and time, and are therefore able to reconcile classical mechanics and the theory of electromagnetism but are very far from common sense.

After the paradigm shift introduced by Einstein, on the one hand, the theoretical concepts of physics are no longer in accordance with the spontaneous intuitions of common sense about the movements of macroscopic bodies but, on the other hand, the concepts of STR and GTR allow you to see how much the conceptualization of the world offered by direct experience and common sense is naïve and deceptive. The change of paradigm Einstein introduced in physics teaches us that if the development of a science gives rise to a conflict between the “scientific image” and the “manifest image” of the world, it is the latter that must be laboriously and counter-intuitively adapted to the former and not vice versa.

### Consciousness naturalized

It might seem a stretch to draw a parallel between the paradigm shift that swept physics a century ago due to the formulation of the theory of relativity and the present cognitive turn in our approach to understanding ourselves driven by research in the cognitive sciences more generally and neuroscience in particular. However, there are at least some similarities between the suspicion and disbelief with which STR and GTR were regarded by many scientists and philosophers a century ago and the objections raised against the possibility of naturalizing consciousness today.

In other words, the resistance to the reduction of consciousness to its neural correlates we see today is similar to the alleged impossibility of substituting Newton’s concept of absolute time with Einstein’s concept of time pronounced a century ago by many philosophers and scientists. For example, H. Bergson the most distinguished French philosopher at the beginning of 20th century even wrote a book against STR to support the difference, in his view unbridgeable, between the “spatialized” time of physics and “duration”, the interior psychological continuous flowing of consciousness. According to Bergson, the former is a simple theoretical construct of science, that is, an instrument of calculation. The latter is instead something that you introspectively and with absolute certainty recognize as real.

If the comparison between these debates inspired by two paradigm shifts in different epochs appears too superficial to be taken seriously, it must be added that the theoretical and formal similarity between the two pa-
radigm shifts is far more profound than a simple analogy between the resistance encountering by both conceptual revolutions. In order to highlight such a deep similarity it is necessary to accompany the brief reconstruction of STR and GTR previously given with an equally synthetic survey of the changes introduced by the cognitive sciences (especially neuroscience) in current philosophical concepts of mind and consciousness.

Of course, if you wish to naturalize the mind and consciousness you cannot unfortunately rely on already fully developed theories like STR and GTR. However, you can freely use the countless studies on the nature of consciousness published by neuroscientists in the last twenty years to present some guidelines for a scientific and naturalistic reconstruction of the concepts of mind and consciousness.

First of all, let me define what is the naturalization of a mental phenomenon. Mental phenomena that are of type X under a certain introspective-phenomenological reconstruction in the language of folk psychology can be naturalized only if the following conditions are met:

(1) Mental phenomena of type X must be “functionally reducible” to functional states of type Y (Y is a concept of cognitive psychology).
(2) The functional states of type Y must be implemented by brain processes (or more generally by biological processes) of type Z (Z is a concept of biology).

According to this definition of naturalization procedures, all mental phenomena can be reduced to brain processes (and more generally to biological and physico-chemical processes) by combining a functional reduction with a neuro-biological implementation. Let us now apply this scheme to the naturalization of the states of consciousness. The first step is to clarify what might be an acceptable introspective-phenomenological reconstruction of what is commonly meant by words such as “consciousness” or “awareness”. Among the concepts offered by the most respected studies of today in the field of cognitive psychology and neurosciences a good candidate to fulfil this task is the concept of Complex Scene proposed by Edelman in the wake of W. James.

In this Jamesian perspective (in some respects reminiscent of the thought of H. Bergson as well) consciousness is private and subjective and is an ongoing flow of perceptions, emotions, feelings, desires, intentions and thoughts that is always continuous and referred to the same subject but at the same time always changing. Consciousness is selective and has distinct contents in distinct moments but at the same time it is always the bearer of a synoptic coherent view of reality.

That said, what is the concept best suited to fulfill the role of functional reconstruction of Edelman’s Complex Scene among the neuro-biological and psychological concepts available today?

Two concepts are strong contenders: the cognitive-psychological concept of Global Workspace Memory and the neurological concept of Dynamic Core. Although different in many respects, both concepts converge in representing consciousness as the characteristic of brain dynamics that allow the mental states implemented by brain processes to exchange information with each other and thereby to issue motor responses reciprocally consistent and effective.

Consequently, it is reasonable to hypothesize that all the neural processes that at a certain moment “enter” (in functional terms) the Global Workspace (or, it is just the same, the Dynamic Core) of a person’s brain thereby become able to exchange with each other the information that they are conveying and converge (in purely functional terms again) into the Complex Scene experienced by that person in that moment. But how can a brain process be admitted, as it were, to the Global Workspace (or to the Dynamic Core)?

Although so far no one has given this question a definitive response it is plausible
to think that only brain processes which have assumed a certain “format” can enter the *Global Workspace* and that this format is literally phenomenal consciousness. In other words, according to this hypothesis phenomenal consciousness is neither an immaterial “thing” (that is, a soul) nor an alleged *Central Processor Unity* of the human brain functionally similar to the C.P.U. of a computer. Phenomenal consciousness is rather the common “format” employed by all the processes in the brain that make up the *Complex Scene* (or the *Dynamic Core*). And this format must be identified in turn with a property shared by their respective dynamics.

In other words, in accordance with the first step of the naturalization procedure sketched above phenomenal consciousness must be redefined in functional terms as the higher order property of brain dynamics that transforms a sequence of detached and independent brain processes into the continuous stream of an agent’s subjective experience. Moreover, brain dynamics can be represented as a trajectory in a vector multidimensional space whereas its first order properties or higher order properties can be represented by first order derivatives or higher order derivatives.

Moreover, if in accordance with the second step of the naturalization procedure you search for a neural implementation of such a functional format, the synchronization of oscillating neural circuits is the most plausible candidate to carry out (at least in part) that role. This hypothesis was initially proposed as underpinning the neurological correlates of consciousness in the case of visual perception but is now highly credited as an explanation for the emergence of many other conscious states, for example various phenomena of “binding”, which explain e.g. how three distinct kinds of sensory inputs (relative to colors, shapes and movements) are fused into the usual conscious perception of an object that moves on a stationary background. According to this hypothesis, the emergence of consciousness is essentially due to the synchronization of brain processes.

Regardless of the value of this hypothesis, what should be emphasized here is the fact that the basic difference between reductionistic and anti-reductionistic theories of consciousness does not rely on neuroscientific technicalities but on the philosophical framework on which such theories are based. Non-reductionists think that no robot could in principle become conscious even if it were possible to equip its artificial brain with all the mechanisms of synchronization (or with any other property of brain dynamics) that reductionists might think responsible for the emergence of consciousness in human beings. Anti-reductionists think so because of a Cartesian conception of the mind according to which mind and body are metaphysically different “things”. Reductionists suggest vice versa that you must not remain prisoner of the image of yourself that you have inherited from philosophical dualism since such an image is falsified by the cognitive and neurological sciences.

When you have acquired a new scientific view of consciousness and you have reached the conviction that consciousness, whatever it may be, is anyway and first of all a neurobiological process among many others, you have already rejected dualism as a solution to the mind-body problem and therefore you have already changed the image that you have of yourself as a human being. After such a conceptual revolution in folk psychology comparable to a paradigm shift in a scientific theory, the identification of consciousness with the synchronization of brain processes or with any other property of brain dynamics ceases to be implausible.

### Consciousness, time and gravity

Is there something in common between the transformation of the concepts of time and gravity brought about by STR and GTR on the one hand and the “cognitive turn” that reduces consciousness to be the common format of the brain processes which enter the *Global Workspace* (or the *Dynamic Core*) on
the other hand? Along with many indubitable differences there are also some interesting similarities between these two cultural revolutions.

First of all, the concept of “inclined time” proposed by STR on the one hand and the neuroscientific concept of consciousness I have previously spoken of as a certain property of brain dynamics on the other hand are both theoretical concepts whose construction depends on a common methodological choice, that is, the use of vectors to represent the respective objects: physical phenomena or psycho-neurological phenomena.

More in detail, in GTR – as previously explained – the force of gravity is absorbed, as it were, by the curvature of a non-Euclidean vector space. However, if you want to give a simplified representation of the motion of a body in the framework of classical mechanics you can manage to do it provided that the body moves very slowly in comparison to the light speed. But you have to pay a price for this choice favouring classical mechanics: in your representation the curvature of space-time disappears and consequently gravitation comes out again as an external force whose origin is mysterious.

Now, let me come back to my hypothesis that the brain dynamics that implements the Global Workspace (or the Dynamic Core) can be meta-represented by means of vectors. If this is true, then it is possible to venture the further hypothesis that the brain monitors itself and acquires a better control of its own activity by constructing a simplified internal model of its own dynamics. Let me admit that this further hypothesis is also true. In this case, would it not be plausible to think that in the construction of a simplified model of its own activity the brain tends to isolate the neural correlates of phenomenal consciousness (that is, the synchronization of certain brain processes and perhaps other features of brain dynamics that are not known so far) from the rest of its own dynamics?

In other words, would it not be plausible to think that a scientific vector meta-representation of the model constructed by the brain of its own activity would show that in this model consciousness (especially with regard to voluntary actions) is represented as an external force able to teleologically drive the brain activity by blocking automatic motor responses? And might it not be the case that this vector meta-representation of the way in which the brain represents consciousness to itself reveals a certain formal similarity with the way in which classical mechanics represents gravity?

This hypothesis would explain why each human being spontaneously tends to identify himself with a conscious and free agent who is the master of one’s own body and not a part of it (or of its activity).

In other words, as classical mechanics must represent gravity as a mysterious external force since it ignores the curvature of space-time in a similar way it might be that the brain in the representation that it gives of its own activity in the “format of consciousness” (which is its internal “machine code”) tends to represent its own dynamics as the activity of a conscious free agent, that is, a Self. As a matter of fact such a Self is only a virtual (and in part fictive) entity implemented by distributed properties of brain dynamics.34 What seems to be a free agent who drives the body is in fact a part of the body and its activity, that is, something embedded and embodied in the brain. Why does the brain deceive itself (that is, us) on this point?

That is not easy to say. But the most likely hypothesis is that biological evolution has given human brains an internal “machine code” capable of simplifying brain activity. The brain often uses “computer tricks” such as the digitalization of the electrical signals that run along axons in order to minimize losses in the amount of information that is transferred from one brain area to another.35

Might it not be that every human being has a spontaneous inclination to think (or better to feel) that he/she is the “sailor” of his/her own body, that is, he/she is a free and self-conscious agent although in fact this is
false? Might it not be that this self-deception was installed by biological evolution in the brain of all human beings because it makes their sensori-motor coordination (and more generally their interaction with the external world) more successful although they are «nothing but a packet of neurons» or better their being conscious and self-conscious is a “servo-mechanism” of their brain? But if their being conscious and self-conscious is a servo-mechanism, then it is a “rebel servo-mechanism”!

In other words, the biological mechanism of consciousness and the Self selected by Mother Nature to improve sensori-motor coordination changed its function step by step through a huge exaptation process. The result of this process is that the brain of every human being after his/her birth is progressively “programmed” by the experience of social life, that is, by learning a language and a culture. Biological evolution selected in humans a brain that is able not only to perform very good sensori-motor coordination but also to live in society.

And so that servo-mechanism of your brain that in every moment creates in you the illusion of being a conscious and free agent if you are awake makes of you a person, too, insofar as it creates in you the further illusion of being a “subject” who on the one hand can maintain her/his personal identity over her/his whole life (despite the plurality and mutability of social roles that you have taken on) and on the other hand interacts with other human beings insofar as you recognize them, too, as persons with a conscious mind like your mind. Thus, each human being becomes in his/her eyes a “person-Self” who in fact is only a fiction produced by the dynamics of his/her brain and a certain socio-cultural context. However, this illusion is necessary for her/his mental health and to execute voluntary acts.

---

### Time, the sense of time and consciousness

To sum up, it is a plausible hypothesis that you appear to yourself as a conscious mind distinct from your body only because this illusion is a trick found by biological evolution to adapt certain animals to their environment. If it were possible to prove that this hypothesis is true, what else would be required to definitively admit that consciousness and the Self are neuro-biological phenomena?

In this case your consciousness and your Self would necessarily appear to you as something ontologically different from the activity of your brain, although this is a mere illusion, because this illusion is created by the brain itself and is biologically useful.

Certainly, to prove that this hypothesis is true is a task that cannot be fulfilled by a philosopher through any arguments a priori. Evidence is needed. However, such evidence is becoming available. For example the German neurologist E. Pöppel has maintained that human beings perceive the flow of time by means of two mechanisms of brain synchronization. The first mechanism creates a “window” of thirty milliseconds: all visual stimuli that fall within this window are perceived as simultaneous. The second window, which lasts about two or three seconds, is constructed by putting together hundred of windows of the first kind: all the visual stimuli that fall into this second window are perceived in sequence but at the same time as a single “extended present” or “specious present”. In other words, the psychological present has a duration of about two or three seconds. Simultaneity and the present are constructions of the brain.

What matters here is that the two brain mechanisms by which we perceive time (simultaneity and the present) are synchronization mechanisms of oscillating brain circuits identical to those that – as explained above – are responsible for the emergence of consciousness. Therefore, the neurosciences confirm one of the key points of Kant’s transcendental aesthetics: the (pure) intuition of time is the form of the internal sense, that is, the common “format” of all mental states. More exactly, the perception of being
awake coincides with the internal sense that time goes by.

Now, if you combine this conclusion which comes from neurobiology with STR you can finally explain why STR inevitably appears counter-intuitive to us human beings: the ratio between the speed of the bodies whose motion is directly observable and the speed of light is too small to produce time dilations and space contractions that can be detected to the naked eye. However, this response becomes misleading if it is not specified that the order of magnitude according to which the ratio between the speed of bodies that are directly observable and the speed of light must be small. Smallness is a comparative concept.

If we consider the problem from this point of view it is clear that the scale at which that relation appears small is a human scale. More precisely, it is the scale at which the human brain works to control sensori-motor coordination. The reason why the absolute time of classical mechanics seems intuitively acceptable whereas the dilatable time of STR is counter-intuitive must not be searched for in the structure of objective external reality but in the way in which the human brain perceives this reality.

And it is here that the theory of brain synchronization I have previously referred to is very useful. According to this theory – as we have seen – all the visual stimuli that fall within a range of 30 ms. are perceived as simultaneous. Consequently, for the human brain (and therefore for all human beings in their everyday life) the speed of light is practically infinite within a radius of 9,000 km (the distance covered by light in 30 ms). But an infinite speed of light makes it so that the Lorentz Transformations coincide with Galilei Transformations and that the space-time of Minkowski is transformed into a vector space in which time-axes of all reference systems are parallel to each other and the units of measurement of time and space are independent of each other. In other words, it is the limited power of temporal resolution of the human brain that makes the absolute time of classical mechanics intuitively appear acceptable whereas the dilatable time of STR is counter-intuitive.

More generally, psycho-neurological theories of perception such as Pöppel’s theory suggest that phenomenal time is not identical to real time. Moreover, if – as we have seen – the perception of time coincides with the emergence of consciousness, then it is plausible to assume that the relation between phenomenal time and real time is analogous to the relation between phenomenal consciousness and its neural correlates. Consequently, if you can find out in the functioning of the brain an explanation for the difference between phenomenal time and real time, why should you not explain by certain properties of your brain’s dynamics why you feel like a free agent and a self-conscious master of your own body whereas you are in reality just a servo-mechanism of your brain?

To sum up, the Hard Problem is neither a true scientific problem that has not been empirically solved yet nor a mystery. It seems to be an insoluble mystery in principle only because it is in fact a philosophical pseudo-problem which is not to be solved but rather dissolved through conceptual analysis. However, the conceptual analysis able to dissolve it (or at least to avoid having it rest in the centre of the image that we human beings have of ourselves) cannot be reduced to an analysis of common language in the manner of L. Wittgenstein. Such an analysis must nourish itself, as it were, with the new ideas coming from the cognitive sciences in general and from cognitive neuroscience in particular. Einstein’s theories constrained physicists to correct and in part abandon some fundamental concepts of classical mechanics. Why should a new science of the mind/brain not deeply change folk psychology?

In other words, the effort that is usually required to accept STR and GTR because of their being strongly counter-intuitive is comparable to the effort that is required to overcome the reluctance of many people in ac-
cepting the fact that we human beings are just a “packet of neurons”. But the advantage that we get in both cases thanks to these efforts is equally big. STR and GTR, particularly if coupled with Pöppel’s neurological theories, are able to explain why such theories are counter-intuitive and less acceptable than classical mechanics for common sense although they are preferable to classical mechanics from a scientific point of view. In a similar way, a new science of the mind/brain allows us to now begin to clarify within the “scientific image of the man” – to borrow Sellars’ words – why the “manifest image” that we human beings have of ourselves usually favours Cartesian dualism but is fundamentally wrong.

### Acknowledgement

I am very thankful to Giuseppe Trautteur for the advice that he gave me after reading a first draft of this paper. I alone am responsible for the bad use that I may have made of his advice.

### Notes


5. Since the speed of light is $c \cong 300,000 \text{ km/s}$, then the speed of the spaceship is $v=0.8c \cong 240,000 \text{ km/s}$.

6. Both for $O$ and for $O'$ (if their clocks always remain synchronized in accordance with the principle of classical mechanics that time is absolute) the signal sent from the spaceship at its departure reaches the station after $4h$. In the same time the spaceship will travel a distance of $3.2lh$ ($4h \times 0.8c = 3.2lh$). Therefore, the return signal must cover only $4lh - 3.2lh = 0.8lh$ and according to $O$ the speed at which the signal should approach the spaceship in its return journey should be $c+v=1.8c$. Consequently the signal should cover the distance of $0.8lh$ in $0.8lh$: $1.8c = 0.8lh$.

7. Since the experiment I am proposing is a thought experiment, all the empirical data resulting from the observations of $O$ and $O'$ are in fact assumptions that are part of the construction of the experiment itself. However, these assumptions are consistent with all empirical data actually available through real experiments.

8. Consequently, the space measures taken by $O$ and $O'$ also must be different since the constancy of light speed in $K$ and $K'$ requires that the ratio between space units of measurement and time units of measurement is $s/t=c=1$ in both reference systems.

9. I shall explain later why according to STR time coordinates must be expressed both in $K$ and in $K'$ not by real numbers (e.g. $t$ or $t'$) but by imaginary numbers multiplied by the real number $c$ that expresses the light speed (e.g. $ict$ or $ict'$).

10. I have assumed in my thought experiment that all events represented in Figure 1 and Figure 2 have space coordinates $y=0$ and $z=0$. If this is not the case then by applying the Pythagorean Theorem the following formula is valid for all events that lie in a world line of light: $(x^2+y^2+z^2) = 0$.

11. I mark ($b^*$) with an asterisk to indicate that this formula will turn out to be wrong according to STR.

12. The price of the exception that you make if you do not change the sign of $\tau$ is that the space-time of STR is not Euclidean any more since the length of $O-E1$ qua hypotenuse of the right-angle triangle whose catheti are $x$ and $t$ is no longer the sum of the squares of the catheti but their difference. However, the advantage is that you avoid the use of imaginary numbers in all cases in which $t$ is bigger than $x$, that is, in all real cases. For according to STR the speed of any body is necessarily less than the light speed.

13. If $c=1$ then $x'=(x-vt)/(1-v^2)^{1/2}$ and $t'=(t-vx)/(1-v^2)^{1/2}$.
To automatically calculate Lorentz Transformations you can use, for example, the animation available at <http://www.trell.org/div/minkowski.html>.

This is the reason why I wrote $t$ and $t'$ instead of $ict$ and $ict'$ in (a)-(d).

«If we wish to describe the motion of a material point, we give the values of its co-ordinates as functions of the time. Now we must bear carefully in mind that a mathematical description of this kind has no physical meaning unless we are quite clear as to what we understand by “time.” We have to take into account that all our judgments in which time plays a part are always judgements of simultaneous events. If, for instance, I say, “That train arrives here at 7 o’clock,” I mean something like this: “The pointing of the small hand of my watch to 7 and the arrival of the train are simultaneous events”» (A. Einstein, On the Electrodynamics of Moving Bodies (1905), in: H.A. Lorentz, A. Einstein, H. Minkowski, H. Weyl, The Principle of Relativity. A Collection of Original Memoirs on the Special and General Theory of Relativity, edited by A. Sommerfeld, W. Perrett, G.B. Jeffrey, Dover Publication, Dover (USA) 1923, pp. 37-65, here p. 39).

This does not mean that the conceptual difference between the three dimensions of space properly understood and the fourth dimension of time is lost in STR. This difference reappears in STR thanks to the fact that real numbers are used for the three spatial coordinates of any event whereas imaginary numbers are used for their time coordinates.

As is well known, as to the cause of gravitation Newton himself wrote his famous sentence: “Hypotheses non fingo” (“I frame no hypotheses”), (see I. Newton, The Mechanical Principles of Natural Philosophy, edited by B. Motte, London 1729, vol. II, p. 392).


See G.M. Edelman, Wider than the Sky, cit.


See G.M. Edelman, Wider than the Sky, cit.


For a vector representation of mental states see P.M. Churchland, The Engine of Reason, the Seat of the Soul, cit. For a general introduction to brain dynamics see M.I. Rabinovich, K.J. Friston (eds), Principles of Brain Dynamics: Global State Interactions (Computational Neuroscience), MIT Press, Cambridge (MA) 2012.


36 F. Crick, *The Astonishing Hypothesis*, cit., p. 3.


