



Human Consciousness: Where Is It From and What Is It for

Boris Kotchoubey*

Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Tübingen, Germany

Consciousness is not a process in the brain but a kind of behavior that, of course, is controlled by the brain like any other behavior. Human consciousness emerges on the interface between three components of animal behavior: communication, play, and the use of tools. These three components interact on the basis of anticipatory behavioral control, which is common for all complex forms of animal life. All three do not exclusively distinguish our close relatives, i.e., primates, but are broadly presented among various species of mammals, birds, and even cephalopods; however, their particular combination in humans is unique. The interaction between communication and play yields symbolic games, most importantly language; the interaction between symbols and tools results in human praxis. Taken together, this gives rise to a mechanism that allows a creature, instead of performing controlling actions overtly, to play forward the corresponding behavioral options in a “second reality” of objectively (by means of tools) grounded symbolic systems. The theory possesses the following properties: (1) It is anti-reductionist and anti-eliminativist, and yet, human consciousness is considered as a purely natural (biological) phenomenon. (2) It avoids epiphenomenalism and indicates in which conditions human consciousness has evolutionary advantages, and in which it may even be disadvantageous. (3) It allows to easily explain the most typical features of consciousness, such as objectivity, seriality and limited resources, the relationship between consciousness and explicit memory, the feeling of conscious agency, etc.

OPEN ACCESS

Edited by:

Lieven Decock,
VU University Amsterdam,
Netherlands

Reviewed by:

Fred Keijzer,
University of Groningen, Netherlands
Julian Kiverstein,
Academic Medical Center (AMC),
Netherlands

*Correspondence:

Boris Kotchoubey
boris.kotchoubey@uni-tuebingen.de

Specialty section:

This article was submitted to
Theoretical and Philosophical
Psychology,
a section of the journal
Frontiers in Psychology

Received: 26 September 2017

Accepted: 04 April 2018

Published: 23 April 2018

Citation:

Kotchoubey B (2018) Human
Consciousness: Where Is It From and
What Is It for. *Front. Psychol.* 9:567.
doi: 10.3389/fpsyg.2018.00567

Keywords: awareness, communication, embodiment, objectivity, play, tool use, virtual reality

The buttock, however, in man, is different from all animals whatsoever. What goes by that name, in other creatures, is only the upper part of the thigh, and by no means similar.

George Louis Leclerc de Buffon (1792, pp. 80–81)

Why do people think? Why do they calculate the thickness of walls of a boiler and do not let the chance determine it? Can a calculated boiler never explode? Of course, it can. We think about actions before we perform them. We make representations of them, but why? We expect and act according the expectancy... Expectancy [is] a preparatory action. It outstretches its arms like a ball player, directs its hands to catch the ball. And the expectancy of a ball player is just that he prepares arms and hands and looks at the ball.

Ludwig Wittgenstein (1996, pp. 109, 139)

The two epigraphs already give partial, but essential, answers to the questions in the title. Where human consciousness is from? In a large extent, it is from the exceptionally extensive tool use, which would be impossible without the erectness supported by the exclusively strong gluteal muscles. What is its function? As indicated by Wittgenstein, it is a set of simulated anticipations.

Notwithstanding substantial differences, most contemporary theories of consciousness (e.g., Dennett, 1991; Damasio, 1999; Edelman and Tononi, 2000; Koch, 2004; Maia and Cleeremans, 2005) regard it as a kind of information *processing*. The present paper, in contrast, regards it as a kind of *behavior*. Behavior is a biological adjustment by means of movements and all kinds of movement-related physiological activity (see Keijzer, 2005, for general principles of the modern theoretical analysis of behavior). Of course, the brain plays a critical role in the control of behavior. Complex forms of behavior (including consciousness) necessarily require, and become possible due to, the complexity of the controlling brain. But there is no isomorphism between a controlling system and a controlled system.

The paper is not about neural correlates of consciousness (NCC). I just do not find the problem of NCC very interesting for several reasons, the simplest of which is: correlation is not causation. Further, it is not about the so called hard problem of consciousness (Chalmers, 1996). The starting point of the present considerations is actively behaving organisms able to various forms of learning (mainly, associative learning). I assume that thus behaving organisms already possess something that can be called “core consciousness” (Damasio, 1999).

By the term “human awareness”, I mean, in accord with most philosophers of mind (e.g., Searle, 2000; Beckermann, 2005), the ability to experience one’s own “internal states” as intentional states (Brentano, 1982), i.e., internal states that are “about” some external objects. This term does *not* imply that all components of this “human awareness” are uniquely human or that this kind of consciousness cannot be found in any nonhuman animal.

Several aspects of the presented model are already described in other published or submitted texts. In such cases only a very brief summary will be given here, and the reader will be referred to further papers. I understand that this way of presentation is highly inconvenient, but the space in open access journals is too valuable to afford the luxury of repetition.

The structure is as follows. First, I describe *precursors* and the three main *behavioral components* giving rise to human consciousness. Second, I describe a “*central block*” of human consciousness built on the interface between these three components (see **Figure 1**). This part is the least original for

the simple reason that description of human consciousness has been undertaken by numerous thinkers from St. Augustin to modern cognitive scientists, and a completely novel description is hardly possible. However, this section is necessary to show how the extant descriptions follow from the three components displayed in the first section, and to put it apart of alternative descriptions.

Third, we describe several most peculiar *features* of human consciousness to show how easily they are deduced from the presented model. Finally, we briefly regard the relationships between this model and some other, similar or remote theories of consciousness. Again, I ask for understanding that, for the above-mentioned space reasons the two latter points cannot be discussed in full in the present text; this discussion remains a topic of a separate analysis.

PRECURSORS

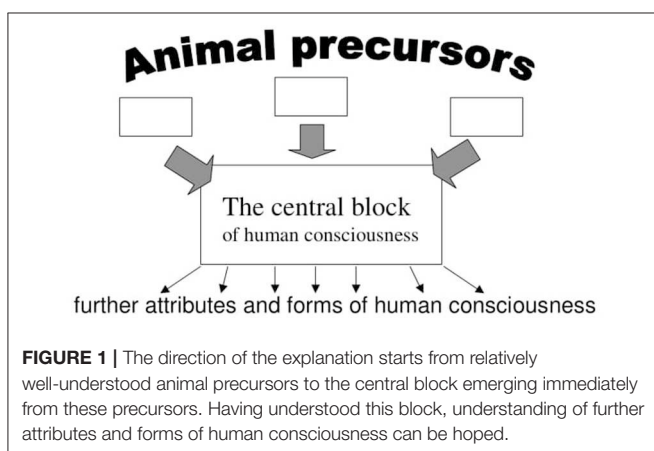
Anticipation, Core Consciousness, and Preconditioning

The organism/environment system is to be kept in a state of extreme energetic nonequilibrium (Schrödinger, 1944). Life, therefore, is the continuous battle against the Second Law of Thermodynamics. All organisms’ needs, from the need of a paramecium in food to the need of a composer to write a symphony, can be subsumed as a need in negentropy, in making order out of energetic death.

To maintain the highly improbable energetic state, organisms interact with their environment in a continuous process of anticipatory regulations. “Regulations” means that environmental disturbances are steadily compensated to make possible the “necessary condition for free an independent life”: the stability of the internal milieu (Bernard, 1865). “Anticipatory” means that physiological regulations at the moment t are such as to compensate the disturbances at the moment $t+1$. This is particularly true for moving animals. The more mobile is an organism, the more distant is the organism’s environment in terms of space, the more ahead of the present point it must be in terms of time.

However, the future cannot be known for sure. Anticipatory processes can therefore be regarded as hypotheses built by the organism about the future state of its environment. All biological adaptations are merely “assumptions to be tested by the selection” (Eibl-Eibesfeldt, 1997, p. 28). Behavioral adaptations, however, are even more hypothetical than, say, morphological adaptations, because they can be tested immediately after an action rather than later in the life. Behavior is principally anticipatory, i.e., based on prediction and correction of upcoming environmental disturbances (Grossberg, 1982; Rosen, 1985; Friston, 2012).

Several authors including, e.g., Bickhard (2005; Bickhard et al., 1989) and Jordan (1998, 2000; Jordan and Ghin, 2006) indicate that anticipatory interactions give rise to core consciousness. The emergence of primary elements of consciousness (“the hard problem”) is beyond the topic of the present article. Of course, it is difficult to know and even more difficult to describe “what



it is like” to have *only* core consciousness. Any description of a qualium (if there is such a thing), requires a term (e.g., “redness”: Humphrey, 2006) which belongs to much higher levels of consciousness than the described phenomenon itself. In any case, our object here is not the emergence of such simple forms of consciousness but a very long way from them to that Cartesian cogito that we usually conceive of as our human awareness.

Anokhin (1974) demonstrated that all mechanisms of conditioning, including both Pavlovian (classical) and Skinnerian (operant) processes, can be considered as anticipatory phenomena. However, an important modification of the classical conditioning procedure is particularly interesting for the following development. In this modification, called preconditioning (Brodgen, 1939), subjects are presented a contingent pair of neutral stimuli (S1–S2; e.g., light and tone), none of them having biological significance. Not surprisingly, their combination does not yield any observable effect. Subsequently, S2 is paired with a typical unconditional stimulus (UCS, e.g., food), leading to a classical conditioned response (CR, salivation). After this, the first neutral stimulus (S1) is presented. Surprisingly, it also elicits a CR, although it has never been combined with the UCS.

The fact that stimuli having no reinforcing value can nevertheless affect behavior was a big challenge for behaviorist theory (see Seidel, 1959, for review). In fact, preconditioning implies two different achievements. First, a new dimension of *time* is opened. The association between S1 and S2 must be retained in memory until S2 is combined with UCS. Even more importantly, the animal’s brain should have sufficient complexity to make use of *contingency of nonsignificant events*. Obviously the animal could not know that one of the stimuli would acquire a meaning in the subsequent conditioning. Therefore, it must be in possession of free resources to record some statistical regularities in the “environmental noise,” whose meaning is momentarily zero. The only purpose of this enormous vast of resources is that some of these presently meaningless combinations may (perhaps) become meaningful in future.

Preconditioning is widely presented among different vertebrate species even in a very young age (e.g., Talk et al., 2002; Barr et al., 2003). Recent data indicate the possibility of sensory preconditioning in bees (Müller et al., 2000) and fruit flies (Brembs, 2002).

Up to this point, the animal lives in the world of its needs, among those relevant events that either satisfy these needs or counteract the satisfaction. This *Lebenswelt* is determined by the genetic design of the animal and continuously extended by learning: new features are revealed as soon as they are associated with genetically hard-wired biologically significant features. An organism is engaged only in those interactions it is designed for (Maturana, 1998). This simple law is first broken by preconditioning: the animal learns to pay attention to *meaningless* events. It overcomes immediate satisfaction of its needs and begins to “save information” with the hope that sometime it may become useful.

Play

Preconditioning is association of two external elements (“stimuli”) having no immediate consequence. Likewise, play is association of *organism’s own actions* and external events, also having no immediate consequence. This definition leads us to distinguish between the immediate and the remote gains of an activity. On the one hand, play is “for fun” (otherwise it should not be called so), on the other hand, it can have a serious function. This vast of energy could not be so ubiquitous in nature if it is not compensated by some important gain. This contradiction between the immediate uselessness and the remote usefulness of play is probably the central point of all ethological, cultural and philosophical discussions to this issue.

Frequently, play appears to copy a serious activity (Burghardt, 2005). Play superficially imitates hunting, sex, aggression, but it is not what it seems. However, the imitative character is not obligatory. When parrots or monkeys just hang on branches and swing without any purpose, most human observers would find that they are playing, although such hanging and swinging does not appear to resemble any serious activity.

Although it is sometimes very difficult to decide whether a particular activity is a play, most people agree that many mammals and birds play (Burghardt, 2005). Playing octopuses have also been described (Kuba et al., 2006). Some groups like predators, primates, sea mammals, and solipeds play distinctly more than others. Further, in all nonhuman species youngsters play considerably longer time (by a factor between 10 and 100) than adults. Already Lorenz (1971) noted that youngsters are more ready to learn nonsense.

An important feature of play is security. In play, skills can be exercised without a risk to fail. When a predator fails in hunting, it can die of hunger. (I mean wild animals, of course, not the pets cared for by old ladies.) A youngster which fails in hunting play will be alimented by its parents. Play, therefore, introduces something that can be called “second reality” (Vygotsky, 1978). In this reality the life is going on as if it is the “primary reality,” but with the nice difference that whenever I don’t like what happens, I simply stop the process and go out, or start it anew. This makes play suspiciously like consciousness. Coaches of athletes are aware of this similarity, thus they combine training without real competition (which is play, as competition is reality) with mental imagery (which is a typical phenomenon of human awareness).

Play is only relatively secure, however. A hunting play takes place in the same real world as the real hunt. Accordingly, although the dualism between play and reality is presented in philosophical thinking (Huizinga, 1950), it is not as strong and troubling as the dualism between mind and matter. Although the result of the playing activity is not of vital importance, the circumstances are real, and obstacles are material, thus the animal can be seriously injured. Remember how many soldiers in the armies die, not in a war, but in training.

Tools

Play is the first important consequence of the ability to learn without reinforcement. The second consequence is the use of tools. The role of tools in creating the world of objects, and of the very distinction between objective and subjective, is analyzed

in Kotchoubey (2014). A tool is a neutral component of the environment used to operate with other components which, in turn, are related to the animal's survival. For example, a stick is neither eatable nor dangerous; but it can be used to reach fruits or to fight enemies. Manipulation with a tool cannot be successful from the very first trial. A stick is eventually manipulated "just for fun," and then, suddenly, it turns out to be useful. Thus no animals unable to play can use tools.

Remember that animals do not worry about the world "as such" (von Uexküll, 1970). They know (i.e., can efficiently deal with: Maturana, 1998) those elements and features of the world which are related to the animal's needs. This can be illustrated by the following scheme:

[Me] < == > [something in the world as related to me]
(note that "me" is presented on both members of the scheme!)

Using a tool, an animal gets knowledge about a new kind of qualities: qualities which relate an element of its environment, not to the animal itself, but to other elements of the environment. For example, a stick used by apes for reaching bananas should possess the feature of hardness:

[Me] < — > [a component of the world] < - - - > [another component related to me]

The dashed line < - - - > indicates a newly acquired relation between two components of the outer world. From this moment I possess a bit of "objective" knowledge, in the sense that I know about some aspects of the world which are unrelated to my own being in this world. Bananas are eatable only as long as there is me who can eat them. Sticks, to the contrary, will remain hard independently of my existence. My knowledge of this hardness is, therefore, objective knowledge, and the corresponding aspect of my environment (i.e., the stick), is now an object. Also the solid line < — > represents a new kind of relations: my ability to operate with tools and to choose means appropriate to my ends.

The objectivity of the obtained knowledge is, however, limited by the fact that the object in question remains related to me in two aspects: as my tool (something in my hand) and as an immediate mean to my goal (the banana). The object is "compressed" between me and me; it does not have an open end. Some animals, however, can make one step more combining two tools, as chimpanzees in famous experiments of Köhler (1926). This can be depicted as:

[Me] < — > [object 1] < ~~~~~ > [object 2] < - - - > [my goal]

The twisting line < ~~~~~ > between two objects stands for a relation from which I am completely factored out. For a long time after these experiments, it was believed that using higher-order tools is limited to most intelligent primates: humans and some apes. Now we know that using and making tools is widely presented among numerous animals including birds. Some parrots (like kea) particularly dexterous in tool usage are also particularly playful (Huber and Gajdon, 2006). Crows, ravens and finches use twigs, petioles, leaf stems and wire in both experimental and natural environment (Hunt et al., 2006; Watanabe and Huber, 2006). They flexibly adjust the length of their tools to the perceived distance from the food, prepare leaf stems for more convenience and use stones to bend wire or make hooks (Weir et al., 2002). They can use a short stick to reach a

long stick to reach the food with the latter (Taylor et al., 2007). Chimpanzees, keas, and New Caledonian crows know something about the *objective world*.

Communication and Symbols

Another important precursor that added to play and tool usage arises from communication. Communication can be defined as a behavior whose *main effect* is changing the behavior of another animal of the same or different species. A similar definition regards communication as the whole of biological effects *selected* for their influence on other animals of the same or different species. Any directed effects on another's behavior require transmission of signals between animals, and signals have a property to stand for something different from the signals themselves. According to the classical tripartition of Peirce, signals can be symbolic, iconic, and indicative (Buchler, 1955). Most animal communication signals belong to the category of indices. Indices are signs causally related to their reference; classical examples are medical symptoms indicating a particular disease. Fever and headache do not just "mean," or refer to, influenza, they are caused by the viral infect. Likewise, a danger scream is not arbitrary (like the word "danger") but physiologically related to the corresponding emotions of fear and anxiety. Also human exclamations such as "Ah!" and "Wow" are causally related to particular emotions and not products of agreement. They are parts of the organism's state and not just signs of this state (Kotchoubey, 2005a,b).

But a danger scream, though physically related to the state of fear, is not physically related to the cause of this state, e.g., the appearance of a predator. Different animal species use completely different signals to signalize the same thing. The signals can be flexibly adjusted to the kind of danger (Donaldson et al., 2007), as vervet monkeys, for example, produce different alarm signals for leopards, eagles, and snakes (Cheney and Seyfarth, 1990). Slowly, the progressive differentiation of signals can yield the ability to integrate them in new ways: a combination of two alarm signals may not be an alarm signal at all, but acquires a different meaning (Zuberbühler, 2002).

Although both animal communication and human speech are mainly vocal, most authors agree that there is probably no continuity between animal cries as indicative signs and human language constituted mainly from symbolic signs (Ackermann et al., 2014); rather, the decisive change was made within gestural communication (Arbib and Rizzolatti, 1996; Corballis, 2002; Arbib and Bota, 2003). Our closest ancestors in the animal world were not the best in vocal communication. Gestures, i.e., signs made by extremities independently of locomotion, exist, however, only in humans and apes and have not been found in any other animal including monkeys (Pollick and de Waal, 2007). Although animal cries also can be modified by context (Flack and de Waal, 2007), gestures are much more context-dependent (Cartmill and Byrne, 2007; Pika and Mitani, 2007), more flexible and thus their causal links to the underlying physiological states are much weaker as compared with vocalizations (Pollick and de Waal, 2007). Gestures thus paved the way for that separation between the signified and the signifying, which is so characteristic for true human language (e.g., Herrmann et al., 2006).

An interaction of communication signals with another factor mentioned above—namely, play—yields an emergent quality of playing *games*. Already Huizinga (1950) noted that any full-blown language is only possible on the basis of play as its precondition in contrast to animal communication that may not be related to play. Even relative independence of signs on their references can be used by an organism possessing the ability to play, which can now play with these signs, recombine them in deliberate combinations. The way is now open for conditional rules and systems of rules which might appear as strict as natural laws but, in contrast to those latter, they are not physically necessary. The notion of language as a symbolic game is frequently attributed to Wittgenstein (e.g., Hacker, 1988, 1990), who repeatedly compared language with chess (Wittgenstein, 2001).

Like the combination of symbols and play produces to the ability to play games, the mutual fertilization of symbols and tools brings about a new set of abilities to put *remote goals* and to relate them to actual sensorimotor interactions. The former complex gives rise to language, the most universal symbolic game all people play; and the latter gives rise to praxis (Frey, 2008). *Language and praxis* are two domains of abilities very specific for humans. Both require a famously unique feature of the human brain, i.e., its strong hemispheric asymmetry; both aphasias and apraxias result almost exclusively from lesions to the left hemisphere. The exact evolutionary process of this interaction is not completely clear. Gibson (1993) proposed that the use of tools may have caused the transition from mainly vocal to primarily gestural communication. On the other hand, Frey's (2008) review indicates that, more plausibly, tool use and gestures first developed as two independent systems and later interacted to produce human practical abilities.

THE SECOND REALITY OF CONSCIOUSNESS

Now we have all the necessary components together. The organism, which already at the beginning could learn through interactions with its environment, has acquired several additional facilities. Being able to manipulate a hierarchy of tools, it can now discriminate objective features of the world, that is, not only the features relating something to the organism but also features relating things to each other. The organism further possesses a developed system of signs and can distinguish (a) between a sign and an internal state, which the sign signifies; as well as (b) between the sign and an external object which the sign refers to. It can play with these signs, recombine them and construct complex systems of arbitrary rules for symbolic games.

Taken separately, communication, play and tool usage are broadly presented in nonhuman animals. None of these abilities can be said to coincide or strongly correlate with thinking or culture. None is limited to one particular group of our human ancestry, e.g., only to primates or mammals. None is related to a particular type of nervous system (Edelman et al., 2005);

in fact, none can be said “a cortical function” because these behaviors are observed in birds (whose telencephalon is only a remote homolog to the mammalian cortex), and sometimes in insects and cephalopods, with their completely different neural morphology.

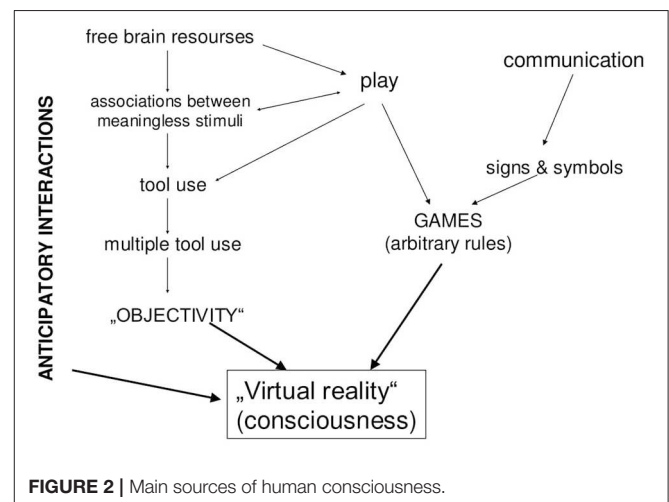
But the *combination* of communicative skills, play and tool use makes a qualitative jump. A being that possesses all these features can do the following: in a difficult situation, in which several behavioral alternatives are possible, it can experience something like “playing forward” (re. Play) each alternative, using symbols (re. Communication) referred to objective knowledge (re. Tools) about its environment. This process, i.e., internalized playing behavioral options, which takes into account objective features of the elements of environment and employs symbolic means, is nothing else but human conscious thinking (Figure 2).

The Virtual Space

The model of human consciousness we use here is a virtual reality (VR) metaphor (Baars, 1998). The main block of human consciousness is anticipatory behavior in a secure “virtual” space of symbolic relationships, in which this behavior does not have any overt consequences. Behavior is thus anticipated by playing it forward in the realm of objectively grounded symbols.

Unfortunately, VR metaphor has also been used in a meaning completely different from the present one. Revonsuo (1995) regarded all experience as VR created by our brains and having nothing in common with “something there.” “The neural mechanisms bringing about any sort of sentience are buried inside our skulls and thus cannot reach out from there—the non-virtual world outside the skull is ... black and imperceptible” (p. 13 of the electronic source). He proposed a horrifying description of reality as “The Black Planet” in which we “cannot see anything, hear anything, feel anything.” We can only construct a virtual world, but the real world outside will forever remain “silent and dark.”

In a similar vein, Metzinger (2003) developed an elaborated model of consciousness largely based on the Revonsuo's (1995) VR metaphor:



“Neither the object component nor the physical body carrying the human brain has to exist, in principle, to phenomenally experience yourself as being related to certain external or virtual objects. ... Any physical structure functionally isomorphic to the minimally sufficient neural correlate of the overall-reality model ... will realize first-person phenomenology. A brain in a vat, of course, could—if approximately stimulated—activate the conscious experience of being a self in attending to the color of the book in its hands, in currently understanding the semantic contents of the sentences being read, or in selecting a particular, phenomenally simulated action ... the discovery of the correlates ... could never help us decide is the question of how the brain in a vat could ever know that it is in this situation—or how you could know that you are now not this brain.” (Metzinger, 2003, p. 415).

This *solipsist* VR metaphor should be strongly distinguished from the *instrumentalist* VR metaphor presented here. By the term “virtual” I mean “artificial” or “second-order,” but *not* “unreal,” “illusory,” or “apparent.” Play hunting is an artificial activity as compared with real hunting, but it is not an illusion of hunting. For example, in Pavlovian second-order conditioning an animal responds to a tone paired with a light after the light has been paired with food. The animal salivates to the tone although it has never been paired immediately with food, but only with the light. Pavlov might have called the light in such experiments a “virtual reinforcer” if the term “virtual” was current those days. But he would have been greatly surprised if a philosopher explained him that what he means is just an “appearance” of reinforcement, that there is no such thing as food in the animal’s world, and that there is no world whatsoever around the animal.

Speaking that human awareness “models” or “simulates” reality, we should understand the meaning of the corresponding terms (Northoff, 2013; van Hemmen, 2014). Every physical or mathematical model building necessarily presumes some pre-knowledge about the process to-be modeled. To build a model of something we must already have an idea of what this thing is and what properties it has. Model building is always a way of testing the hypotheses formulated before we have started to model. Who states that our conscious awareness (or the brain as its organ) models the world, presumes that the world already exists independently of our awareness (Kotchoubey et al., 2016).

The obvious advantage of virtual behavior is the ability of secure learning. The price for learning can be an error and the price for an error can be a failure, an injury or even death. Thus the optimum would be an area in which we may commit errors without being punished for them but, nevertheless, learning from them. This virtual area is consciousness. In words of Karl Popper, instead of dying as a result of our errors, we can let our hypotheses die on our site (Popper, 1963).

On the other hand, the adaptive action is postponed (vast of time), and the resources are consumed for virtual activity having no immediate effect (vast of energy). Therefore, conscious behavior is worth only in particularly complex situations in which its gains outweigh its losses: (1) when there are several behavioral options whose consequences are unclear, or whose chances appear similar; (2) when a risk of negative consequences of a wrong action is high, i.e., when the situation is dangerous but the danger is not immediate. Then, the disadvantage of the delay

is overbalanced by the advantage of withholding an erroneous action.

Therefore, speaking about “human awareness,” we do not mean that humans are in this state all the time. This would be a catastrophic vast of resources that the Mother Nature never could afford. Again, this does not mean that otherwise we behave “unconsciously.” We just interact with our environment, we are *engaged* in our world. As Heidegger (1963) showed, this engagement is beyond the dichotomy of conscious vs. unconscious. We live in this engagement and experience it, and in this sense, we are conscious (Kadar and Effken, 1994), but we do not consciously think. We are just there (Heidegger, 1963; Clark, 1997).

This “ordinary state” of human existence can, of course, be compared with animal consciousness. We are thrown in our *Lebenswelt* like animals are thrown in theirs. However, our world is not theirs. The world in which we are engaged even without exerting conscious control is a social and instrumental world, i.e., it is already built up of those elements (tools, communication, symbolic games) which gave rise our conscious behavior. Many important differences between humans and apes result from the differences in their environment (Boesch, 2006). Most of our activity is automatic, like in animals, but these automatisms differ in both design and content (e.g., Bargh and Chartrand, 1999). Our existential reality is cultural, and so are our automatisms. Whitehead, who claimed that “civilization advances by extending the number of operations which we can perform without thinking” (Whitehead, 1911, p. 61), illustrated this idea not with automatic catching or grasping, but with automatized operations of experienced mathematicians over algebraic symbols. This automatization is hardly attainable even for very clever apes.

Objections

Three arguments, which are frequently put forward to defend the solipsist variety of VR, can be regarded as objections against the present instrumentalist version: illusions, dreams, and paralysis. The illusion argument reads that consciousness cannot be deduced from interactions between organism and reality because in some mental states (e.g., illusions or delusions) we experience something different from reality. The argument is, first, inconsistent because who says that illusory perception is different from reality implies that he knows reality. Second, the argument confuses adaptive and maladaptive consciousness. Illusions in humans adapted to their environment occur only in very specific conditions on the background of millions of cases of veridical perception. Those whose mental life is prevailed by illusions, in contrast, are usually unable to survive without modern medicine and social support. The argument misses, therefore, the adaptive function of awareness.

The use of dreams in this context is equally misleading. The neural basis of dreams, the REM sleep, is completely different from the waking state in terms of physiological, humoral, neurochemical and behavioral components (Hobson and Pace-Schott, 2002; Pace-Schott and Hobson, 2002). Accordingly, dream experience has a number of formal (regardless of dream content) properties qualitatively different from those of our ordinary experience (e.g., Hobson, 1988; Stickgold et al., 2009).

Lisiewski (2006) distinguishes between a “strong” and “weak” VR in the set-theoretical sense. “Strong” VR keeps constraints close to the constraints in the real world as it is experienced in simple forms of sentience. Examples are typical existing VR programs. In a “weak” VR, in contrast, all constraints are removed, e.g., one can fly, be simultaneously two persons, observe oneself from the side, etc. Dream consciousness, like science fiction stories, belongs to the latter category. That is why in dreams the muscle tone is nil, a finding predicted by Freud half a century before this fact was empirically proven (Freud, 1953). Freud’s reason was that when reality constraints are removed, subjects must not have a possibility to actively behave. Therefore, dreams cannot be used as a model of conscious experience because the very essence of dream states is the blockade of the interaction between the organism and its environment.

The paralysis arguments indicates that humans extremely paralyzed for a long time (locked-in syndrome: LiS) and, therefore, lacking all overt behavior, nevertheless retain consciousness and can demonstrate very high intellectual functions (e.g., Kotchoubey et al., 2003; Birbaumer et al., 2008). However, all described cases of LiS (mainly a result of a brainstem stroke or of severe neurodegenerative diseases: for a review see Kotchoubey and Lotze, 2013) concern adult individuals, usually older than 30. All of them possess many years of experience of active behavior. From a philosophical viewpoint it might be intriguing whether consciousness would develop in a child with an inborn LiS, but from the ethical viewpoint we should be glad that no such case is known.

In most LiS patients, at least some movements (usually, vertical eye movements) still remain intact. Due to the progress of medicine and assistive technology, now many locked-in patients can use the remaining movements for communication and describe their experience in the acute period of the disease (e.g., Bauby, 1997; Pantke, 2017). These patients’ reports show that the patients, albeit conscious and in possession of higher cognitive abilities, do not have “experiences as usual” as was previously believed. Rather, LiS is related to subtle disorders of conscious perception and attention that cannot be explained by the lesion of motor pathways but probably result from the dropout of motor feedback (Kübler and Kotchoubey, 2007; Kotchoubey and Lotze, 2013).

Using a Brain-Computer-Interface (BCI), cognitive abilities of LiS and other severely paralyzed patients can be checked independently of their motor impairment (e.g., Birbaumer et al., 2008). The analysis of the corresponding data shows that the ability to learn in patients, who possess minimal remaining movements (maybe only one), is only slightly, if at all, impaired in comparison with healthy controls. However, as soon as this last motor ability is lost, the learning ability is completely lost too (Kübler and Birbaumer, 2008). According to these authors, the minimum capacity to interact with the environment and to be reinforced for successive actions is a necessary prerequisite of intentional learning.

The paralysis argument is also important because it helps to distinguish between phylogenetic, ontogenetic, and actual genetic aspects of consciousness. My main claim that human consciousness emerges in human evolution at an interface

between play, tool use, and communication does not imply that the same three components necessarily participate, and in the same constellation, in the individual development of conscious thinking in children. Even less is it to say that the same components necessarily participate in each actual case of conscious thinking. In the development of human behavior, many feedback loops, originally running between the organism and the environment or between the brain and bodily periphery, later on become shorter and remain within the simulating brain (e.g., Grusz, 2004).

To summarize, the objections are not convincing. The former two miss the adaptive function of consciousness and, therefore, presume epiphenomenalism. The paralysis argument requires an additional assumption of similarity between phylo- and ontogenesis and, besides this, ignores the fact that in the rare documented cases of complete LiS (i.e., when no behavior exists), no cognitive function could be found.

Recursion

The most important implications of the model are recurrent processes at several levels. The best studied class of them is symbolic recursion. When we possess symbols causally free from the elements of the environment they stand for, and arbitrary rules which govern the use of these symbols, we can build symbols which stand for symbols, and rules governing other rules, and symbols standing for a system of rules governing symbols, etc., etc. At least primates (e.g., Flack and de Waal, 2007) and possibly dogs (Watanabe and Huber, 2006) are already capable of metacommunication, i.e., to signals indicating how the following signals should be interpreted. According to Chomsky (1968, 1981), symbolic recursion builds the basis for the infinite complexity of human language.

The second class encompasses instrumental recursive processes automatically emerging with the increasing order of tools. They can constitute highly complex loops, e.g., a machine A produces the machine B which produces the machine C which produces screws necessary to produce the machines A, B, and C.

Recursivity of human consciousness allows us to rebuff another objection that is traditionally put forward against all kinds of instrumentalism. According to the argument, we just need not see objects as tools. For example, we consciously perceive a meadow and trees upon it. Although all of this can be used (cows may pasture on the meadow; fruits may rife on the trees), this usefulness is not normally presented in our consciousness while we are looking at this view. Even in the instance of obviously useful objects (e.g., furniture) we usually perceive these objects without being aware of their instrumental value. I just see my sofa without being afforded to sit down on it. Even less am I aware of continuous action preparation in my consciousness. I see this, hear that, like one thing and dislike another one, but I do not plan any actions with them.

This “neutrality illusion” can be easily explained on the basis of the present model. While each tool use diminishes our personal relatedness to objects, closed loops can completely extinguish this relatedness. In the world of recursive relations between tools producing tools for making tools, I do not feel my personal involvement any longer, but remain an external

observer, a passive viewer of the complex systemic relationships between the multiple components of my world, rather than a component of this same system.

Perhaps the most interesting kind of recursive processes, in addition to symbolic and instrumental recursion, is the *temporal* recursion. It will be briefly depicted below in section Memory.

PROPERTIES

A theoretical model of consciousness can be evaluated by the easiness with which the known properties can be deduced from it (Seth et al., 2005). These authors proposed a list of “sixteen widely recognized properties of consciousness” including philosophical (e.g., self-attribution), psychological (e.g., facilitated learning), and physiological (e.g., thalamocortical loops) criteria. Some of these criteria correspond to our everyday experience (“consciousness is useful for voluntary decision making”), whereas others (“sensory binding”) are only valid within the framework of a particular hypothesis which is neither empirically supported (Fotheringham and Young, 1998; Riesenhuber and Poggio, 1999; Vanderwolf, 2000) nor logically consistent (Bennett and Hacker, 2003; van Duijn and Bem, 2005). The list does *not* include intentionality (Brentano, 1982), nor does it warrant that the 16 criteria are mutually independent (e.g., “involvement of the thalamocortical core” and “widespread brain effects”) and non-contradictory (e.g., “fleetingness” and “stability”).

As our present topic is *human* consciousness, this list of properties to test the model should not, on the one hand, include unspecific properties of any kind of conscious experience. On the other hand, we should not consider properties related to specific historical and cultural forms of human consciousness (e.g., *satori*), and by no way should we regard those questionable “properties” deduced from particular views on human nature, e.g., such property of consciousness as “computational complexity” (as if the WWW were computationally simple).

On the basis of these considerations, and taking into account space limits, I do not claim to check an exhaustive list of criterial properties of human consciousness, but rather, to illustrate the consequences of the hypothesis using a few representative examples: seriality and limited capacity; objectivity; the intimate relation between consciousness and memory; and the sense of conscious agency.

Seriality

The serial character of human conscious experience is a highly salient and, from the point of view of many neurophysiologists, an almost mysterious feature. While the brain (which is supposed to be the seat of mind) works as a parallel distributed network with virtually unlimited resources, conscious events are consecutive, happen one at a moment, and their momentary capacity is strongly limited. Theories regarding consciousness as a particular case of brain information processing must, therefore, suggest a specific mechanism creating serial processing from parallel. This compromises the aesthetics of the corresponding theories, because in addition to the multiple

brain mechanisms generating conscious contents one more mechanism is postulated to make these contents run in a row.

The difficulties disappear, however, we assume that consciousness has been emerged from behavior and is itself a covert behavior. As already said, human consciousness can be afforded only in specific, particularly complex situations. But any kind of complex behavior is a *series* of organism-environment interactions. A cat do not hunt and wash, or eat and play, at the same time. Likewise, we cannot simultaneously turn left and turn right, notwithstanding all parallel distributed processing in our brain.

An example of locomotion, which is largely unconscious, illustrates the limits of parallel behavior. With automatization of a motor skill organisms acquire the ability to perform some motor acts simultaneously. This process plays a particular role in actively communicating animals such as primates. After extensive experience the muscles of face and tongue become independent of peripheral coordinations, and we can walk and talk at the same time. But as soon as the situation gets more complex, this ability to perform two behavioral acts in parallel is lost. It is difficult and even dangerous to actively communicate while descending an unfamiliar stair in complete darkness. Complex behaviors are serial by nature. In those exceptional cases in which they can run in parallel, the states of consciousness can be parallel, too: whales sleep with one half of the body.

A similar idea was suggested by Merker (2007) who related the seriality of conscious behavior to the existence of a “motor bottleneck” in the bridge (*pons cerebri*) in the upper part of brain stem. However distributed are processes in the cortex, in order to reach muscles the cortical activity must pass through the site where all impulses from the forebrain to the motor periphery converge. Locked-in syndrome discussed in the section Objections above is most frequently a result of a stroke in this area. From this point of view, consciousness is serial because it is restricted by the “common final path” to the effectors, and its limited capacity is a function of the limited capacity of motor activity.

The serial character of human consciousness is closely related to another specific feature, the intolerance of contradiction. Parallel distributed processing in brain networks has nothing against contradiction; different neuronal assemblies can simultaneously process different aspects of information, perhaps incompatible with each other. Both Freudian and cognitive unconscious (Shevrin and Dickman, 1980; Kihlstrom, 1987) are highly tolerant against contradiction. This fact strongly contradicts (sorry for word play) to the negative affect we get as soon as we notice a contradiction between two contents of our consciousness. Again, the paradox disappears when we realize that consciousness is not a processing but a behavior. Mostly, ambiguous behavior is either physically impossible (e.g., simultaneous moving in two directions), or maladaptive (e.g., making one step forth and one step back). Why should consciousness be different?

Objectivity

This term is used in two interrelated meanings. First, it means that we live in the world which appears to contain distinct

and relatively stable entities called objects. Second, “objectivity” means a kind of detachment, i.e., freedom from values, needs, passions, and particular interests. To my knowledge these features are either taken for granted (i.e., the world just consists of objects), or attributed solely to the development of language (e.g., Maturana and Varela, 1990; Dennett, 1991; Damasio, 1999). The former view is not tenable: our being in the world is not a cold and impersonal observation. The latter view is partially true. Operations with signs standing for something different increase the distance between us and the world. Symbolic systems are powerful tools we use to deconstruct the complex world into separate things.

However, in order to use words as tools, we first must use tools. Language can support but not create the objective world. Only tools can do this because they are material components of the same world. Tools put themselves between us and our needs projected into the world (see Tools above). They expand the space relating the organism to its immediate *Lebenswelt* so much that they transform it into the space separating the organism from its environment. They enforce me to deal with relationships between different elements of the world, and between different features of these elements, rather than to be ever egocentrically busy by the relationships between the world and myself. They decentralize world. More than one million years ago, the early Homo already employed higher-order tools (Ambrose, 2001; Carbonell et al., 2008). Long before Copernicus stopped the sun rotating around the Earth, tool usage stopped the world to rotate around each animal’s needs. In the extent the needs retire into the background, so the related emotions. We can now be engaged into the world of entities which do not immediately concern us. We can, within some narrow limits, remain cool and “objective” (Kotchoubey, 2014).

Symbolic games add a new quality to this objectivity. The two sets of recursive loops (the symbolic and the instrumental) mutually interact, further enhancing detachment and disengagement. When the recursivity of tools added with the recursivity of signs conditionally referred to the tools, the distance between the organism and the world becomes a gap. First the relationship “me-world” was replaced, or at least complemented, with the relationships among objects. Then even these latter are substituted by the relationships between arbitrary symbols standing for objects and their relations.

The higher is the order of tool use, the stronger am I bracketed out of the chain of events. The transformation of the fluent energy of the world into the static order of stable objects finally attains the degree at which I am myself (almost) similar to other things. The living human organism, which is primarily a node of struggling, suffering, enjoying, wanting energies, becomes (almost) just another object of cold cognition among many objects. In the course of this decentration an individual can even get a strange idea that his/her own behavior is caused by external objects, like a behavior of a billiard ball is mechanically caused by other balls hitting it!

In our culture, the objectivity of the world is further strengthened and enhanced by the stance of natural science (Galilei, 1982). From this point of view only quantitative relations among elements of the world are “real,” that is, they belong to

the world itself. In contrast, qualities, i.e., the effects of these relations on my own organism, are regarded not as “physical facts” anymore, but as “illusions” of my “mind” (Dewey, cf. Hickman, 1998). Thus color, warmth, loudness, all these proofs of my engagement in the world are just “mental events” indirectly referred to some other (real, physical) characteristics such as wavelength, molecular energy, or amplitudes of oscillations. Interpretation of the relationships *between* us and various aspects of our environment in terms of the relationships *among* these aspects became a criterion of scientific truth.

Thus the physiological opposition between the *milieu intérieure* and *milieu extérieure* (Bernard, 1865) becomes the philosophical opposition between the Subject and the Object. Both are products of using tools, separating the organism from the world.

Memory

The relationship between embodiment, memory, and consciousness are discussed in a parallel paper (Kotchoubey, in press) and can only briefly be concerned here. Human consciousness defined as a virtual space for covert anticipatory actions implies an ability to deliberately delay reinforcement (“building a bridge over the gap of time”), thus introducing a strong time dimension. It has even been proposed that the freedom to operate in time, i.e., to activate in one’s behavior temporary remote events, conditions, and one’s own states, is the *specificum humanum*, the main feature distinguishing humans from all other creatures (Bischof-Köhler, 2000; Suddendorf, 2013). The close correspondence between kinds of memory and kinds of consciousness was first demonstrated by Tulving (1985a,b). Also, he showed on the basis of neuropsychological observations that memory is a bi-directional function, i.e., it relates the organism not only with its past but also with its future (Tulving, 1985b, 1987).

In accordance with this idea, the present model of consciousness is hardly compatible with the classical view that memory is about the past. This view is based on the computer metaphor implying strong separation between storing and processing, which does not exist in biological organisms. From an evolutionary viewpoint, memory was selected not to store the past but to use the past in order to optimize the present behavior and to organize future adaptation. “[T]here can be no selective advantage to mentally reconstruct the past unless it matters for the present or future” (Suddendorf and Busby, 2005, p. 111). This is equally true for short-time memory (STM) as an obvious derivative from *working memory*, which is immediate future planning (e.g., Atkinson et al., 1974). Atkinson and Shiffrin (1971) even identified the actual content of consciousness with the content of STM.

As soon as memory is not regarded anymore as a function of saving information, but rather, as that of behavioral adaptation taking into account the organism’s past, many phenomena usually viewed as “errors of memory” became normal. When we are prompted to remember something, we build hypotheses, check them up and adjust them according to the actual situation to other components of knowledge (Bartlett, 1967) as well as to social conditions (Loftus, 1997; Gabbert et al., 2003). In other words, our behavior toward the past does not differ from

that toward the present or future. Most so-called errors of memory are not malfunctions, they indicate the flexibility and adaptability of our behavior in the time domain (Erdelyi, 2006). Remembering is neither a faithful recapitulation of past events nor a construction of a reality-independent mental world, but interaction and adaptation (Suddendorf and Busby, 2003).

In the VR of human consciousness an overt action with its real consequences is delayed until the virtual action is virtually rewarded or punished. Therefore, the time dimension, which originally was a flow of events, is now split into several axes. First, the flow of behavioral events is held, as long as no events happen. Second, the sequence of events in consciousness creates a new flow of symbolic events: a second axis, along which we can free move in both directions (Bischof-Köhler, 2000; Suddendorf, 2013). The freedom of moving backwards is of vital importance; otherwise, erroneous actions with their negative consequences would be as uncorrectable as they are in real life. Third, although overt behavior is delayed, other processes (physiological activity at the cellular, tissue and organ levels, as well as automatic actions) go on.

The split time makes human consciousness particularly interesting and dramatic. The combination of the resting external time with the free travel in the virtual time provides us with the ability to quickly actualize (in the sense: make actual, efficient in our behavior) any remote or possible consequence. If we only once ask, “when I do this, what after?” nothing (in principle) can prevent us from asking the same question infinitely many times (e.g., “when after X comes Y, what comes after Y? And what is after the event which will happen after Y?; etc.). This recursive process renders us to know that the day after tomorrow will be followed by the day after the day after tomorrow, and so on up to the day of our death. But, then, what happens after my death? I don’t want to say that all humans really ask these endless “what after?” questions. I want to stress, however, that the ability to realize one’s whole life and death and to ask what will follow it is not a product of a particular cultural development, but belongs to the most universal properties of human consciousness and immediately results from the basic structure of anticipatory behavior in the virtual space of symbolic games.

Voluntary Action

The question, why complex human (and animal) behavior is necessarily free, has been discussed in many details elsewhere (Kotchoubey, 2010, 2012). In this text, I shall only concern one particular aspect of this general problem, namely the strong *feeling* of agency, of personal control of one’s actions. This issue clearly demonstrates the advantage of the present model of human consciousness over the prevailing cognitive models. These latter assume that the brain first has to make representations of outer objects, and then, this cognitive activity is complemented by actions to deal with these objects. Despite a century of serious critique (Dewey, 1896; Järvillehto, 2001b) this notion is still alive and leads us to ask the question of how the brain can know that my movements belong to me. As always, the answer is postulating an additional “module” of attribution of actions to the agent (de Vignemont and Fourneret, 2004). Thus a cat’s brain first makes a decision that she will *jump* for

a mouse, and then, she needs an additional decision making that *she* (rather than another cat, or a fox, or a raven) will jump for the mouse.

Such problems do not emerge altogether when we remember that the object of adaptive behavioral control are not our motor actions (the output) but a particular state of affairs (the input) (Marken, 1988; Jordan, 1999). Humans think in teleological terms (Csibra and Gergely, 2007) not because such thinking can be useful but because actions cannot be described in terms other than their outcomes (Hommel et al., 2001). Actions are voluntary if the input patterns they generate can be covertly tested within the virtual space of consciousness.

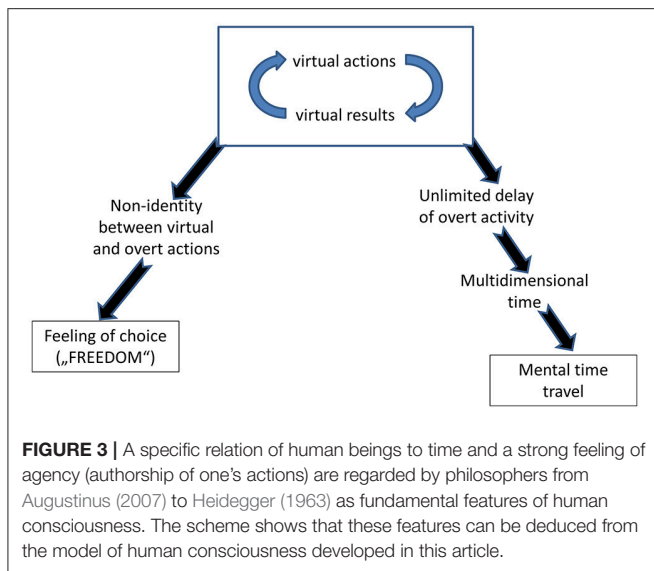
This definition has important corollaries. It does not require that we are aware of any details of the actions we nevertheless perceive as conscious. The logical impossibility of such awareness was demonstrated by Levy (2005). Equally impossible (and in a blatant contradiction with our intuitive feeling) are the demands that voluntary actions should be preceded by feelings like “wish” or “urge,” or must imply a zero effect of the situation on behavior (e.g., Wegner, 2002). No behavior can be carried out without taking some aspects of the environment into account.

The basis of agency is the simple fact that predators, as a rule, do not attack their own body. This is the difference between “the inside” and “the outside” quite similar to the distinction between the own and alien albumins in the immune system. Of course, this fundamental representation of behavioral actions as “mine” need not be conscious, let alone conscious in the sense of the present article. However, as soon as we admit that consciousness develops from behavior, we understand that this simple me/non-me distinction is a precursor of human agency.

What makes this agency the fact of our conscious awareness is the choice. Most lay people simply identify freedom with choice (e.g., Monroe and Malle, 2010; Vonasch and Baumeister, 2013; Beran, 2015). Choice is the result of the fact that virtually performed actions can differ from the actions overtly performed. If there is no this difference, i.e., if we always perform the same action that we have thought about, the whole enterprise of “consciousness as VR” would be meaningless. But when this difference exists, it proves that in the same situation at least two different actions were possible, and therefore, we had freedom of choice (Figure 3). In hindsight, we regard an action as voluntary if we did, or could, estimate possible consequences of several alternatives and selected one whose virtual results were the best¹.

A necessary but not sufficient mechanism of this choice is inhibition of overt behavior. Therefore, the view that associates volition with the ability to exert inhibitory control of otherwise involuntary actions (veto: Libet, 1985) deserves attention. Human conscious activity strongly correlates with activation of those brain structures whose main function is inhibition. These structures are specifically active during particularly complex forms of human behavior. However, inhibitory control is a precondition of volition but not volition itself. If I have to repair

¹An in-depth discussion about the meaningfulness of the question, whether a different action could be chosen in the case of the exact repetition of the same situation, is given in Kotchoubey (2012, 2017). To put it briefly, meaningless is primarily the term “exact repetition.”



my car, I must stop it first; but stopping is not repair. The decisive point is not veto but choice.

OTHER VIEWS

A good theory does not only shed light at its object, but also at the other views on the same object. As a famous example, the relativity theory not just explains the mechanisms of the Universe; it is also successful in the explanation of why other respectable theories (e.g., Newtonian mechanics) gave a different explanation of the same mechanisms: because they regarded a very limited range of possible velocities. Likewise, from the point of view presented here the origins of several alternative theories of consciousness can be apprehended. Of course, this highly interesting task cannot be pursued in full in this paper; we cannot discuss all existent theories of consciousness in their relationship with the current model. Rather, I shall restrict the review to the approaches apparently similar to the present one.

Embodiment Theories

The proposed theory is most similar to embodiment theories of consciousness, simply because it is one of them. Embodiment theories are characterized by “4Es”: human experience is *embodied* (i.e., brain activity can realize mental processes only being involved in closed loops with the peripheral nervous system and the whole bodily periphery including inner organs and the motor apparatus), *embedded* (i.e., the brain-body system is further involved in closed loops with the environment), *enacted* (i.e., in the interaction with the environment, brain and mind not just process information but make the organism to play the active role *pursuing its goals* notwithstanding environmental disturbances), and *extended* (i.e., it involves external devices as if they were parts of the same brain-body system) (e.g., Tschacher and Bergomi, 2011).

Beyond the general agreement at these four points, different embodiment theories of mind and consciousness build a very

broad spectrum varying in their account on the exact role and mechanisms of realization of each point, as well as interactions between them. The hard discussions running in the last decades within the embodiment camp would, however, lead us far beyond our present topic; they are addressed, e.g., in the publications of Menary (2010), Bickhard (2016), Stapleton (2016), Tewes (2016), and the literature cited there.

To be sure, the present approach shares these four E-points. Particularly, anticipatory regulations we have begun with, are closely related to the principles of embeddedness and enactiveness; and the critical role of tools in my approach fully corresponds to the principle of extendedness.

However, to my best knowledge no embodiment theory has up to date been devoted to the issue of the origin and the biological basis of specifically human awareness. Rather, several representatives of this approach attacked the hard problem of the origin of elementary forms of sentience or perceptual experience (e.g., Varela et al., 1991; O'Regan and Noe, 2001; Jordan, 2003; Bickhard, 2005; Noe, 2005; Jordan and Ghin, 2006). How successful these attacks have been, should be discussed elsewhere. From my point of view, the sensorimotor theory (Hurley, 1998; Noe, 2005) has not convincingly responded to arguments raised by its critics (e.g., Hardcastle, 2001; Kurthen, 2001; Oberauer, 2001) who indicated that even a best explanation of mechanisms and phenomena of perception does not imply an explanation of perceptual experience, that is, “what it is like” to perceive a red color or a high-pitch tone. If we assume that simple robots do not have conscious experience, the fact that the proposed embodied and enacted mechanisms of perception can be modeled in robots already refutes the idea that these mechanisms can explain consciousness.

The sensorimotor theory is, of course, only one of the embodiment-grounded attempts to explain the emergence of consciousness. Other (“interactivist”: Bickhard, 2005, 2016); or (“relational”: Jordan and Ghin, 2006) approaches, having a more profound evolutionary foundation, may be more successful in this enterprise. Nevertheless, they have not yet given any systematic account of the transition from the alleged simple sentience to human consciousness, which is the theme of the present paper.

Simulation Theories

Part 2 above exposed the idea that human consciousness is a secure space where behavioral actions are virtually performed, and their consequences are virtually apprehended. In general, this idea is not new but goes back to the British associationism of the eighteenth century (Hesslow, 2002). In experimental psychology, the concept of cognition as “covert trials” was advanced by Tolman (e.g., Tolman and Gleitman, 1949; Tolman and Postman, 1954) and in philosophy, as the theory of conjectures and refutations (Popper, 1963). It is further in line with the well-known scheme of “test-operation-test-exit” (Miller et al., 1960). About 40 years ago, Ingvar (1979; also Ingvar and Philipsson, 1977) practically formulated the concept of consciousness as anticipatory simulations; unfortunately, he justified his conclusions by brain imaging data which appear to be

of questionable quality today, not surprising given the enormous progress of brain imaging techniques since then.

The same idea of covert behavior underlies the concept of efference copy (von Holst and Mittelstaedt, 1950), as well as some control-theoretical models that regard themselves as alternatives to the efference copy theory (e.g., Hershberger, 1998). In the last decades similar views were thoroughly developed under the terms “simulation” (Hesslow, 2002, 2012) and “emulation” (Grusz, 2004). Particularly interesting from the present viewpoint are the data that virtual performance of actions includes anticipation of action results with simultaneous inhibition of the overt execution of these actions (Hesslow, 2012). Behavior, originally realized in large feedforward loops including bodily periphery and the environment, can subsequently be reduced to the loops within the brain.

Notwithstanding the clear similarity between my VR metaphor and all these old and recent views, there are substantial differences as well. Thus the notion of motor simulation frequently defines “behavior” as purely motor activity separated from perception and anticipation of results. The present approach is, in contrast, based on the presumption of the control theory that behavior is *control of input* rather than control of output and cannot, therefore, be regarded as a set of commands sent to muscles. The very sense of a virtual behavior is obtaining its virtual consequences. A related minor point is the idea shared by many adepts of simulated behavior that a motor system has some “point of no return,” so that when a simulated motor activity attains this point, the movement cannot be inhibited anymore and must be executed. The concept of the “point of no return” is a leftover of motor control ideas of the nineteenth century and has no place in the modern movement science (e.g., Latash, 2008, 2012).

But notwithstanding these rather minor differences between all these approaches (regarded above in a cavalry manner) and the present one, there is a very big difference in the *kind of explanation*. The primary interest of simulation theorists is a *how-explanation*. They ask, how, i.e., using what kind of mechanisms, virtual behavior is realized. My point, to the contrary, is a *why-explanation*: why virtual behavior is realized thus and not differently. For example, without the phylogenetic roots in playing behavior, simulated activity could not possess its astonishing freedom to initiate any virtual action in any circumstances, to interrupt or terminate at any deliberate point and to re-start at any moment. The components of communication and tool usage also have profound effects on the nature of human consciousness, as we shall see in the next session.

Language and Thought

The matter of this section is not properly a theory or a class of theories but rather a group of loosely related views, converging on a literal understanding of “con-sciousness” as “shared knowledge” (lat con-scientia). Thus consciousness is regarded as the product of cognitive activity converted into a form of language to be shared with others. The idea that, roughly speaking, “human consciousness is animal consciousness multiplied by language” has, in fact, become a matter of general consensus as a component of almost all theories of consciousness,

even between the authors so different as Edelman (1989), Dennett (1991), Maturana (1998), Järvillehto (2001a,b), and Humphrey (2006), who barely agree at any other point. Indeed, what else is specific for human (in contrast to the animal) consciousness if not the fact that it is based on social cooperation and language-mediated communication?

Crudely, many socio-linguistic theories may be classified into pre-structuralist (e.g., Vygotsky), structuralist (e.g., Levi-Strauss) and post-structuralist (e.g., Derrida). The first stress the process of internalization in which social (interpersonal) processes are transformed into internal cognitive (intrapersonal) processes. Consciousness, from this point of view, is the pattern of social relations (for example, a child-parent interaction) transported into the head. The second class of theories contends that consciousness is based upon hidden cultural and linguistic stereotypes (e.g., the opposition “cultural” vs. “natural”) creating stable, largely a-historic structures. The third view insists on the virtually absolute relativity of the structure and content of conscious human behavior and (in contrast to structuralism) its historical and ideological interpenetration.

Above, when discussing the interaction of communication and play, we already mentioned that human consciousness is frequently regarded as a symbolic game, and that this view is usually traced back to Wittgenstein: “*The limits of my language mean the limits of my world*” (Wittgenstein, 1963; Section 5.6, emphasis in original).

This view, however, leaves unclear wherefrom the structures (or the rules of the game) take their stability and causal power if they are not filled by the content of a language-independent world. Post-structuralists capitalized on this inconsequence and proposed a radical solution for the above problem: if consciousness does not have any meaningful content besides the rules and structures of the game, then, it does not have any rules and structures either (Derrida, 1975). Thus even the notion of symbolic game became much too restrictive since it may imply that there is something the symbols stand for—but in fact, they stand for nothing. Any kind of human behavior is just a “text,” which can be interpreted in a variety of possible ways. For itself (i.e., before and beyond this process of interpretation) there is no such thing as the meaning of an action. Also the world, the so-called “nature” or “reality,” is a text to be interpreted and deconstructed (Foucault, 1966). Not only, therefore, everything is merely a sequence of signs, but these signs do not signify anything: the classical opposition between the signifying and the signified (de Saussure, 1983) is thus annulled. Hence, consciousness is not a game, as previous socio-linguistic theories regarded it, but rather a free play (Derrida, 1975) whose rules may appear and disappear like clouds in a windy day. From the early socio-linguistic point of view, consciousness is its own manifestation in systems of signs. From the later socio-linguistic point of view, consciousness is just these systems of signs and nothing more. “Cognition is a relationship of language to language” (Foucault, 1966; Ch. 1.4).

One can say that these views evolved from the theories of socio-linguistic *foundation* of consciousness, peaking in the linguistic *determinism* in Wittgenstein (1963) and Whorf (1962), to the theories of the unlimited *freedom* of consciousness in its historic and linguistic realization. This freedom, from their

(and my) point of view, largely roots in the freedom of the sign, which, in its development from index to symbol, abandoned its causal link to its reference. Importantly, the notion of language as a symbolic game is not limited by syntax. Rather, it is the very meaning of the words which is determined by their location within the network of tacit verbal rules. E.g., we cannot understand the meaning of the word “hard” without its oppositions such as “soft” or “easy.” Also the meaning of mental concepts is nothing but their usage in language, i.e., their position in the linguistic game. Understanding consciousness means understanding how the term “consciousness” is used in our culture (Bennett and Hacker, 2003).

Because many very influential linguistic theories originally accrued in philology and cultural anthropology, they may appear to concern only particular forms of consciousness expressed, e.g., in arts and letters but not the basics of human consciousness. This is not true. These ideas profoundly affected the contemporary thinking on mind and consciousness down to the level of such “elementary” functions as visual perception (e.g., Gregory, 1988, 1997) and neural development (Mareshal et al., 2007). They left their trace even in strongly biological approaches to cognition and consciousness (e.g., Varela et al., 1991; Maturana, 1998).

From the present point of view, socio-linguistic theories correctly emphasize communication and play as important sources of human consciousness. Most elaborated of them also stress its prospective nature making conscious behavior “free” in the sense of being not determined by the past. However, all these views, traditional and contemporary, philosophically or biologically oriented, completely miss the instrumental nature of human behavior. Many of them talk about tools; e.g., they regard words as tools, scientific theories as tools, etc. But besides this, our consciousness is based on simply tools, which are not words, not theories, just tools. Using them, we either attain our goal (if we correctly grasp the objective relation between elements of the environment and their properties), or not (if our conceptions are wrong). Thus the results of tool usage continuously test the validity of our symbolic games. “By their fruit you will recognize them” (Bible: Mt. 7, 16). This fruit is the banana, which Köhler’s (1926) apes reached using sticks and boxes. If their concepts of sticks and boxes were true, they reached the banana, but when they were false, they remained hungry.

It is true that, e.g., a building can be regarded as a “text,” and that the architect may have projected his personality into his drawings. But in addition, the building has to withstand gravity, wind and possibly earthquakes. To understand the meaning of “hardness,” it is important to recognize its relationships within the field of its use in the language (e.g., the hardware/software distinction). But it is also important to remember that our ancestors failed to reach bananas using a bundle of straw, simply because the bundle was not hard.

Common Working Space

The theory of common working space (CWS: Baars, 1988, 1997) is probably the most elaborated *psychological* theory of consciousness in the last 30 years. The theory regards the mind as a coordinated activity of numerous highly specialized cognitive modules (Fodor, 1981, 1983) whose work is largely automatic. When some of these specialists meet a processing task for which

no preprogrammed solution is available, they build a CWS to make this task as well as all proposed solutions open for every other module. This can be compared with a large audience in which many small groups work each with its own problem, but there is also a possibility to broadcast a problem for the whole audience. Consciousness is this broadcasting; there is a competition for access to it, because the space is only one, and the tasks are many. Therefore, the most interesting processes determining the content of our consciousness are not those which happen in consciousness but those which decide what specialized module(s) should get access to it.

The CWS theory not only provides an explanation for very many characteristic properties of consciousness, but it is also quite compatible with other interesting theories (e.g., reentrance theory), which we cannot discuss due to space limitation.

The metaphor of consciousness behind the CWS model is that of a theater (Baars, 1997). The CWS can be regarded as an open scene accessible for all cognitive modules. The similarity between the theater metaphor and the VR metaphor is obvious. Both presume a scenery, a show, thus pointing to one of the key components of the present hypothesis, i.e., *play*. Both theater and VR are spaces where things are played.

But in this play, we should not play down the differences between the two metaphors. A theater presumes many spectators, who rather *passively observe* the actors’ activity, whereas a VR is concentrated around a single participant, who is *actively engaged* in this reality. Furthermore, arbitrariness is much stronger in the theater than in the VR. Millions of people admire opera theater in which they witness how personages express their emotions by continuous singing, which would appear strange and silly in real life.

Also interestingly, the theater metaphor does not warrant the uniqueness of consciousness. Many cities have several theaters, and some people can visit two or three on an evening. Nevertheless, the most established version of the CWS theory assumed that there exists only one common space for each brain (and each body, Shanahan, 2005). Many concrete predictions of the CWS theory result from the assumption of the strong *competition* between modules striving for the access to the only possibility to broadcast. Later on Baars (2003) suggested that there can be multiple CWSs working in parallel. This raises questions such as: what can count for a space to be regarded as “common,” and how many specialized processors (may be only two?) should be connected to build a “partial consciousness.”

It cannot be denied that we normally experience one particular state of consciousness each moment, in accord with the old philosophical idea of the “unity of consciousness” (James, 1890). Baars (1997) and Dennett (1991) devoted a lot of intriguing pages to the issue of how this unity can be created by the distributed brain. Neuroscientists (Singer, 1999; Treisman, 1999; Tallon-Baudry, 2004) regard this question as the main question of the neurophysiological underpinnings of consciousness.

Thus we are surprised that we have only one state of consciousness at one time, despite millions of parallel functioning neuronal circuits in our brain. However, we are not surprised when a big animal (e.g., a whale) makes a jump as a whole, although its body consists of many thousands simultaneously (and to a large extent, independently) working cells. We don’t

regard this unity as a miracle and don't postulate a specific mechanism of binding these cells into a single organism.

Complex behavior is realized in the form of muscular synergies (Bernstein, 1967; Gelfand et al., 1971; Turvey, 1990; Latash, 2008), which dominate the actual distribution of muscle forces each moment of time. These synergies are motor equivalents of the CWS. The unity of consciousness is the unity of behavior. This does not mean that the unity is unproblematic, but the analogy with motor control indicates the correct name for the problem. The motor system does not have a binding problem but must solve the problem of excessive degrees of freedom, also called "Bernstein problem" (Bernstein, 1967; Requin et al., 1984; Latash, 2012). The principle of "freezing degrees of freedom" implies that muscles are not permitted to work independently, but all must remain within a frame of a unifying synergy. With the development of a motor skill, the synergy becomes more and more local until it is limited to those muscles only, whose participation is indispensable.

Of course, when we talk about muscles we also mean the whole nervous apparatus these muscles are connected with. Therefore, as far as the unity of the CWS is the unity of complex behavior, there is no contradiction between the CWS theory and the present one. Accordingly, the control of new, unskilled actions is frequently conscious. The question is *why* the common working place of consciousness is common. From my point of view, it is not because a group of processing modules has decided, in a democratic or dictatorial way, that a given piece of information is interesting enough to make it accessible for the whole audience, but because complex behavior cannot be organized other than by coordinating all activity to a common pattern. Likewise, we do not make two conscious decisions simultaneously not because the two must compete for one scene, but because, if we did make them simultaneously, how would we realize these decisions? The answer is: serially, one after the other.

CONCLUSION

The model is presented that conceives of human consciousness as a product of a phylogenetic interaction of three particular forms

of animal behavior: play, tool use, and communication. When the three components meet in humans, they strengthen and mutually reinforce each other producing positive feedback loop. Therefore, although all three elements of human consciousness are present in many animal species (not necessarily human predecessors), there is no other species that plays, communicates and uses tools as much as humans do.

The suggested three-component structure permits to easily explain most typical features of human conscious awareness: its recursive character, seriality, objectivity, close relation to semantic and episodic memory, etc. Other specific features of human consciousness (e.g., the emotion of anxiety) remain, unfortunately, not discussed due to space limits. Finally, a comparison of the current approach with other theories of consciousness (embodiment theories, simulation theories, common working place) reveals, notwithstanding some similarities, important differences from all of them. Again due to space limits, the complex relationships of this model of consciousness with the multiple draft theory, the re-entrance theory, and the classical dualistic approach must remain outside the present text.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

ACKNOWLEDGMENTS

Fragments of the text were written with the support of the Deutsche Forschungsgemeinschaft (DFG), other portions were supported by Alexander von Humboldt Stiftung. Some ideas presented here emerged in discussions with J. Scott Jordan (Normal, IL) and Timo Järvillehto (Oulu, Finland). Vanessa Singh commented an early version of the manuscript. I am particularly indebted to my students because the general view presented in this paper could only be developed within the interaction during teaching.

REFERENCES

- Ackermann, H., Hage, S. R., and Ziegler, W. (2014). Brain mechanisms of acoustic communication in humans and nonhuman primates: an evolutionary perspective. *Behav. Brain Sci.* 37, 529–604. doi: 10.1017/S0140525X13003099
- Ambrose, S. H. (2001). Paleolithic technology and human evolution. *Science* 291, 1748–1753. doi: 10.1126/science.1059487
- Anokhin, P. K. (1974). *Biology and Neurophysiology of the Conditioned Reflex and Its Role in Adaptive Behavior*, Transl. by R. Dartau, J. Epp, and V. Kirilcuk. Oxford: Pergamon Press.
- Arbib, M. A., and Bota, M. (2003). Language evolution: neural homologies and neuroinformatics. *Neural Netw.* 16, 1237–1260. doi: 10.1016/j.neunet.2003.08.002
- Arbib, M. A., and Rizzolatti, G. (1996). Neural expectations: a possible evolutionary path from manual skills to language. *Commun. Cogn.* 29, 393–424.
- Atkinson, R. C., Hermann, D. J., and Wescourt, K. T. (1974). "Search processes in recognition memory," in *Theories in Cognitive Psychology*, ed R. L. Solso (Potomac, MD: Erlbaum), 101–1456.
- Atkinson, R. C., and Shiffrin, R. M. (1971). The control of short-term memory. *Sci. Am.* 225, 82–90. doi: 10.1038/scientificamerican0871-82
- Augustinus, A. (2007). *Bekenntnisse (Confessiones)*, Transl. by W. Thimme. Düsseldorf: Artemis & Winkler.
- Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Baars, B. J. (1997). *In the Theater of Consciousness*. New York, NY; Oxford: Oxford University Press.
- Baars, B. J. (1998). Metaphors of consciousness and attention in the brain. *Trends Neurosci.* 18, 58–62. doi: 10.1016/S0166-2236(97)01171-5
- Baars, B. J. (2003). "How does a serial, integrated, and very limited stream of consciousness emerge from a nervous system that is mostly unconscious, distributed, parallel, and of enormous capacity?" in *Essential Sources in the Scientific Study of Consciousness*, eds B. J. Baars, W. P. Banks, and J. B. Newman (Cambridge, MA: MIT Press), 1123–1128.
- Bargh, J. A., and Chartrand, T. L. (1999). The unbearable automaticity of being. *Am. Psychol.* 54, 462–479. doi: 10.1037/0003-066X.54.7.462

- Barr, R., Marrott, H., and Rovee-Collier, C. (2003). The role of sensory preconditioning in memory retrieval by preverbal infants. *Learn. Behav.* 31, 111–123. doi: 10.3758/BF03195974
- Bartlett, F. C. (1967). *Remembering: A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press.
- Bauby, J.-D. (1997). *Le Scaphandre et le Papillon, 1st Edn*. Paris: Editions Robert Laffont.
- Beckermann, A. (2005). *Analytische Einführung in die Philosophie des Geistes, 2nd Edn*. Berlin: de Gruyter.
- Bennett, M. R., and Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. Oxford: Blackwell.
- Beran, M. J. (2015). The comparative science of “Self-Control”: what we are talking about? *Front. Psychol.* 6:51. doi: 10.3389/fpsyg.2015.00051
- Bernard, C. (1865). *An Introduction to the Study of Experimental Medicine*. Transl. by H. C. Greene. Henry Schuman, Inc.
- Bernstein, N. A. (1967). *The Coordination and Regulation of Movements*. Oxford: Pergamon.
- Bickhard, M. H. (2005). Consciousness and reflective consciousness. *Philos. Psychol.* 18, 205–218. doi: 10.1080/09515080500169306
- Bickhard, M. H. (2016). Inter- and en-activism: some thoughts and comparisons. *New Ideas Psychol.* 41, 23–32. doi: 10.1016/j.newideapsych.2015.12.002
- Bickhard, M. H., Campell, R. L., and Watson, T. J. (1989). Interactivism and genetic epistemology. *Arch. Psychol.* 57, 99–121.
- Birbaumer, N., Murguialday, A. R., and Cohen, L. (2008). Brain-computer interface in paralysis. *Curr. Opin. Neurol.* 21, 634–638. doi: 10.1097/WCO.0b013e328315ee2d
- Bischof-Köhler, D. (2000). *Kinder auf Zeitreise*. Bern: Huber.
- Boesch, C. (2006). What makes us human (*Homo sapiens*)? The challenge of cognitive cross-species comparison. *J. Comp. Psychol.* 120, 227–240. doi: 10.1037/0735-7036.121.3.227
- Brembs, B. (2002). *An Analysis of Associative Learning in Drosophila at the Flight Simulator*. Würzburg: University of Würzburg.
- Brentano, F. (1982). *Deskriptive Psychologie*. Hamburg: Meiner. 1st publication 1884.
- Brodgen, W. J. (1939). Sensory preconditioning. *J. Exp. Psychol.* 25, 323–332. doi: 10.1037/h0058944
- Buchler, J. (1955). *Philosophical Writings of Peirce*. New York, NY: Dover.
- Buffon, G. L. L. (1792). “A description of man,” in *Buffon’s Natural History, Containing a Theory of the Earth, a General History of Man, of the Brute Creation and of Vegetables, Minerals, etc.* ed J. S Barr (London: Covent Garden), 62–94.
- Burghardt, G. M. (2005). *The Genesis of Animal Play: Testing the Limits*. Boston, MA: MIT Press.
- Carbonell, E., de Castro, J. M. B., Parés, J. M., Pérez-González, A., Cuenca-Bescós, G., Ollé, A., et al. (2008). The first hominin in Europa. *Nature* 452, 465–470. doi: 10.1038/nature06815
- Cartmill, E. A., and Byrne, R. W. (2007). Orangutans modify their gestural signaling according to their audience’s comprehension. *Curr. Biol.* 17, 1345–1348. doi: 10.1016/j.cub.2007.06.069
- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York, NY: Oxford University Press.
- Cheney, D. L., and Seyfarth, R. M. (1990). *How Monkeys See the World*. Chicago, IL: University of Chicago Press.
- Chomsky, N. (1968). Language and the mind. *Psychol. Today* 1, 48–51.
- Chomsky, N. (1981). Knowledge of language: its elements and origins. *Philos. Trans. R. Soc. Lond.* 295, 223–234. doi: 10.1098/rstb.1981.0135
- Clark, A. (1997). *Being There: Bringing Brain, Body and World Together Again*. Cambridge, MA: MIT Press.
- Corballis, M. C. (2002). *From Hand to Mouth: The Origins of Language*. Princeton, NJ: Princeton University Press.
- Csibra, G., and Gergely, G. (2007). “Obsessed with goals”: functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychol.* 124, 60–78. doi: 10.1016/j.actpsy.2006.09.007
- Damasio, A. R. (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. San Diego, CA: Harcourt.
- Dennett, D. (1991). *Consciousness Explained*. Boston, MA; Toronto, ON; London: Little, Brown and Co.
- Derrida, J. (1975). “Structure, sign, and play in the discourse of the human sciences,” in *Jacques Derrida: Writing and Difference* (London: Routledge), 278–294.
- de Saussure, F. (1983). *Course in General Linguistics*. Trans. By R. Harris. La Salle, IL: Open Court.
- de Vignemont, F., and Fournere, P. (2004). The sense of agency: a philosophical and empirical review of the “who” system. *Conscious. Cogn.* 13, 1–19. doi: 10.1016/S1053-8100(03)00022-9
- Dewey, J. (1896). The reflex arc concept in psychology. *Psychol. Rev.* 3, 357–370. doi: 10.1037/h0070405
- Donaldson, M. C., Lachmann, M., and Bergstrom, C. T. (2007). The evolution of functionally referential meaning in a structured world. *J. Theor. Biol.* 246, 225–233. doi: 10.1016/j.jtbi.2006.12.031
- Edelman, D. B., Baars, B. J., and Seth, A. K. (2005). Identifying hallmarks of consciousness in non-mammalian species. *Conscious. Cogn.* 14, 99–118. doi: 10.1016/j.concog.2004.09.001
- Edelman, G. M. (1989). *The Remembered Present: A Biological Theory of Consciousness*. New York, NY: Basic Books, Inc.
- Edelman, G. M., and Tononi, G. (2000). *A Universe of Consciousness. How Matter Becomes Imagination*. New York, NY: Basic Books/Perseus.
- Eibl-Eibesfeldt, I. (1997). *Die Biologie des Menschlichen Verhaltens*. Weyarn: Seehamer.
- Erdelyi, M. H. (2006). The unified theory of repression. *Behav. Brain Sci.* 29, 499–511. doi: 10.1017/S0140525X06009113
- Flack, J. C., and de Waal, F. (2007). Context modulates signal meaning in primate communication. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1581–1586. doi: 10.1073/pnas.0603565104
- Fodor, J. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.
- Fodor, J. A. (1981). *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Cambridge, MA: MIT Press.
- Fotheringham, D. K., and Young, M. P. (1998). “Neural coding schemes for sensory representation: theoretical proposals and empirical evidence,” in *Cognitive Neuroscience*, ed M. D. Rugg (Hove: Psychology Press), 47–76.
- Foucault, M. (1966). *Les Mots et les Choses*. Paris: Gallimard.
- Freud, S. (1953). *The Interpretations of Dreams*, Vol. 4/5. London: Hogarth.
- Frey, S. H. (2008). Tool use, communicative gesture and cerebral asymmetries in the modern human brain. *Philos. Trans. R. Soc. Lond. Ser. B* 363, 1951–1957. doi: 10.1098/rstb.2008.0008
- Friston, K. (2012). Prediction, perception and agency. *Int. J. Psychophysiol.* 83, 248–252. doi: 10.1016/j.ijpsycho.2011.11.014
- Gabbert, F., Memon, A., and Allan, K. (2003). Memory conformity: can eyewitnesses influence each other’s memories for an event? *Appl. Cogn. Psychol.* 17, 533–543. doi: 10.1002/acp.885
- Galilei, G. (1982). *Dialog Über die Beiden Hauptsächlichen Weltsysteme, das Ptolemäische und das Kopernikanische*. Transl. by E. Strauß. Stuttgart: Sxl & von Meyenn.
- Gelfand, I. M., Gurfinkel, V. S., Fomin, S. V., and Tsetlin, M. L. (eds.). (1971). *Models of the Structural-Functional Organization of Certain Biological Systems*. Cambridge, MA: MIT Press.
- Gibson, K. (1993). The evolution of lateral asymmetries, language, tool-use, and intellect. *Am. J. Phys. Anthropol.*, 92, 123–124. doi: 10.1002/ajpa.1330920112
- Gregory, R. L. (1988). “Consciousness in science and philosophy: conscience and con-science,” in *Consciousness in Contemporary Science*, eds A. J. Marcel and E. Bisiach (Oxford: Clarendon Press), 257–272.
- Gregory, R. L. (1997). *Eye and Brain: The Psychology of Seeing*. Princeton: Princeton University Press.
- Grossberg, S. (1982). *Studies in Mind and Brain, Vol. 70*. Dordrecht: Reidel Publishing Company.
- Grusz, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behav. Brain Sci.* 27, 377–442. doi: 10.1017/S0140525X04000093
- Hacker, P. M. S. (1988). Language rules and pseudo-rules. *Lang. Commun.* 8, 159–172. doi: 10.1016/0271-5309(88)90014-6
- Hacker, P. M. S. (1990). Chomsky’s problems. *Lang. Commun.* 10, 127–148. doi: 10.1016/0271-5309(90)90030-F
- Hardcastle, V. G. (2001). Visual perception is not visual awareness. *Behav. Brain Sci.* 24, 985–986. doi: 10.1017/S0140525X98461752
- Heidegger, M. (1963). *Sein und Zeit*. Tübingen: Max Niemeyer.
- Herrmann, E., Melis, A. P., and Tomasello, M. (2006). Apes’ use of iconic cues in the object-choice task. *Anim. Cogn.* 9, 118–130. doi: 10.1007/s10071-005-0013-4

- Hershberger, W. A. (1998). "Control systems with a priori intentions register environmental disturbances a posteriori," in *Systems Theories and A Priori Aspects of Perception*, Vol. 126, ed J. S. Jordan. (Amsterdam: Elsevier), 3–24.
- Hesslow, G. (2002). Conscious thought as simulation of behavior and perception. *Trends Cogn. Sci.* 6, 242–247. doi: 10.1016/S1364-6613(02)01913-7
- Hesslow, G. (2012). The current status of the simulation theory of cognition. *Brain Res.* 1428, 71–79. doi: 10.1016/j.brainres.2011.06.026
- Hickman, L. A. (1998). *The Essential Dewey Vol. 2*. Bloomington: Indiana University Press.
- Hobson, J. A. (1988). *The Dreaming Brain*. New York, NY: Basic Books.
- Hobson, J. A., and Pace-Schott, E. F. (2002). The cognitive neuroscience of sleep: neuronal systems, consciousness and learning. *Nat. Rev. Neurosci.* 3, 679–693. doi: 10.1038/nrn915
- Hommel, B., Müsseler, J., Ascherleben, G., and Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–937. doi: 10.1017/S0140525X01000103
- Huber, L., and Gajdon, G. K. (2006). Technical intelligence in animals: the kea model. *Anim. Cogn.* 9, 295–305. doi: 10.1007/s10071-006-0033-8
- Huizinga, J. (1950). *Homo Ludens: A Study of the Play-Element in Culture*. New York, NY: Roy Publishers.
- Humphrey, N. (2006). *Seeing Red: A Study of Consciousness*. Cambridge, MA: Harvard University Press.
- Hunt, G. R., Rutledge, R. B., and Gray, R. D. (2006). The right tool for the job: what strategies do wild New Caledonian crows use? *Anim. Cogn.* 9, 307–316. doi: 10.1007/s10071-006-0047-2
- Hurley, S. (1998). *Consciousness in Action*. Cambridge, MA: Harvard University Press.
- Ingvær, D. H. (1979). "Hyperfrontal" distribution of the general grey matter flow in resting wakefulness: on the functional anatomy of the conscious state. *Acta Neurol. Scand.* 60, 12–25. doi: 10.1111/j.1600-0404.1979.tb02947.x
- Ingvær, D. H., and Philipsson, L. (1977). Distribution of the cerebral blood flow in the dominant hemisphere during motor ideation and motor performance. *Ann. Neurol.* 2, 230–237. doi: 10.1002/ana.410020309
- James, W. (1890). *Principles of Psychology*. Philadelphia, PA: Henry Holt.
- Järvillehto, T. (2001a). Feeling as knowing - Part II. emotion, consciousness, and brain activity. *Conscious. Emotion* 2, 75–102. doi: 10.1075/ce.2.1.04jar
- Järvillehto, T. (2001b). Consciousness 'within' or 'without'? Review of J. Scott Jordan (ed.), 'modeling consciousness across the disciplines'. *J. Conscious. Stud.* 8, 89–93.
- Jordan, J. S. (1998). "Intentionally, perception, and autocatalytic closure: a potential means of repaying the psychology's conceptual debt," in *System Theories and A Priori Aspects of Perception*, Vol. 126, ed J. S. Jordan (Amsterdam: Elsevier), 181–208.
- Jordan, J. S. (1999). "Cognition and spatial perception: production of output or control of input?" in *Cognitive Contributions to the Perception of Spatial and Temporal Events*, eds G. Ascherleben, T. Bachmann, and J. Müsseler (Amsterdam: Elsevier), 69–90.
- Jordan, J. S. (2000). The role of "control" in an embodied cognition. *Philos. Psychol.* 13, 233–237. doi: 10.1080/09515080050075717
- Jordan, J. S. (2003). Emergence of self and other in perception and action: an event-control approach. *Conscious. Cogn.* 12, 633–646. doi: 10.1016/S1053-8100(03)00075-8
- Jordan, J. S., and Ghin, M. (2006). (Proto-) consciousness as a contextually emergent property of self-sustaining systems. *Mind Matter* 4, 45–68.
- Kadar, E. E., and Effken, J. (1994). Heideggerian meditations on an alternative ontology for ecological psychology: a response to Turvey's proposal. *Ecol. Psychol.* 6, 297–341. doi: 10.1207/s15326969eco0604_4
- Keijzer, F. (2005). Theoretical behaviorism meets embodied cognition: two theoretical analyses of behavior. *Philos. Psychol.* 18, 123–143. doi: 10.1080/09515080500085460
- Kihlstrom, J. F. (1987). The cognitive unconscious. *Science* 237, 1445–1452. doi: 10.1126/science.3629249
- Koch, C. (2004). *The Quest for Consciousness: A Neurobiological Approach*. Englewood, CO: Roberts & Co.
- Köhler, W. (1926). *The Mentality of Apes*. New York, NY: Harcourt, Brace & Co.
- Kotchoubey, B. (2005a). Pragmatics, prosody, and evolution: language is more than a symbolic system. *Behav. Brain Sci.* 28, 136–137. doi: 10.1017/S0140525X05340039
- Kotchoubey, B. (2005b). Seeing and talking: whorf wouldn't be satisfied. *Behav. Brain Sci.* 28, 502–503. doi: 10.1017/S0140525X05360080
- Kotchoubey, B. (2010). Embodied freedom and the escape from uncertainty. *Psyche* 16, 99–107.
- Kotchoubey, B. (2012). *Why Are You Free?* Hauppauge, NY: Nova Science Publishers.
- Kotchoubey, B. (2014). Objectivity of human consciousness is a product of tool usage. *Front. Psychol.* 5:1152. doi: 10.3389/fpsyg.2014.01152
- Kotchoubey, B. (2017). "No chance? – Das Schwellenmodell und die Verschiedenartigkeit des Ungewissen," in *Zufall in der belebten Natur*, ed U. Herkenrath (Hennef: Roman Kovar), 234–257.
- Kotchoubey, B. (in press). A grammar for the mind: embodied and disembodied memory. *J. Conscious. Stud.*
- Kotchoubey, B., Lang, S., Bostanov, V., and Birbaumer, N. (2003). Cortical processing in Guillain-Barre syndrome after years of total immobility. *J. Neurol.* 250, 1121–1123. doi: 10.1007/s00415-003-0132-2
- Kotchoubey, B., and Lotze, M. (2013). Instrumental methods in the diagnostics of locked-in syndrome. *Restor. Neurol. Neurosci.* 31, 25–40. doi: 10.3233/RNN-120249.
- Kotchoubey, B., Tretter, F., Braun, H. A., Buchheim, T., Draguhn, A., Fuchs, T., et al. (2016). Methodological problems on the way to integrative human neuroscience. *Front. Integr. Neurosci.* 10:41. doi: 10.3389/fint.2016.00041
- Kuba, M. J., Byrne, R. A., Meisel, D. V., and Marther, J. A. (2006). When do octopuses play? Effects of repeated testing, object type, age, and food deprivation on object play in *Octopus vulgaris*. *J. Comp. Psychol.* 120, 184–190. doi: 10.1037/0735-7036.120.3.184
- Kübler, A., and Birbaumer, N. (2008). Brain-computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients? *Clin. Neurophysiol.* 119, 2658–2666. doi: 10.1016/j.clinph.2008.06.019
- Kübler, A., and Kotchoubey, B. (2007). Brain-computer interfaces in the continuum of consciousness. *Curr. Opin. Neurol.* 20, 643–649. doi: 10.1097/WCO.0b013e3282f14782
- Kurthen, M. (2001). Consciousness as action: the eliminativist sirens are calling. *Behav. Brain Sci.* 24, 990–991. doi: 10.1017/S0140525X01410119
- Latash, M. L. (2008). *Synergy*. New York, NY: Oxford University Press.
- Latash, M. L. (2012). *Fundamentals of Motor Control*. New York, NY: Academic Press.
- Levy, N. (2005). Libet's impossible demand. *J. Conscious. Stud.* 12, 67–76.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behav. Brain Sci.* 8, 529–566. doi: 10.1017/S0140525X00044903
- Lisiewski, A. M. (2006). The concept of strong and weak virtual reality. *Mind Machine* 16, 201–219. doi: 10.1007/s11023-006-9037-z
- Loftus, E. F. (1997). Creating false memories. *Sci. Am.* 277, 70–75. doi: 10.1038/scientificamerican0997-70
- Lorenz, K. (1971). *Studies in Animal and Human Behavior*. London: Methuen.
- Maia, T. V., and Cleeremans, A. (2005). Consciousness: converging insights from connectionist modeling and neuroscience. *Trends Cogn. Sci.* 9 397–404. doi: 10.1016/j.tics.2005.06.016
- Mareshal, D., Johnson, M. J., Sirois, S., Spratling, M. W., Thomas, M. C. S., and Westermann, G. (2007). *Neuroconstructivism: How the Brain Constructs Cognition*, Vol. 1. New York, NY: Oxford University Press.
- Marken, R. S. (1988). The nature of behavior: control as fact and theory. *Behav. Sci.* 33, 197–205. doi: 10.1002/bs.3830330304
- Maturana, H. R. (1998). *Biologie der Realität*. Frankfurt am Main: Suhrkamp.
- Maturana, H. R., and Varela, F. J. (1990). *Der Baum Der Erkenntnis*. München: Goldmann.
- Menary, R. (ed.). (2010). *The Extended Mind*. Cambridge, MA: MIT Press.
- Merkel, B. (2007). Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. *Behav. Brain Sci.* 30, 63–81. doi: 10.1017/S0140525X07000891
- Metzinger, T. (2003). *Being No One. The Self-Model Theory of Subjectivity*. Cambridge MA: MIT Press.
- Miller, G. A., Galanter, E., and Pribram, K. H. (1960). *Plans and the Structure of Behavior*. New York, NY: Holt, Rinehart and Winston.
- Monroe, A. E., and Malle, B. F. (2010). From uncaused will to conscious choice: the need to study, not speculate about people's folk concept of the free will. *Rev. Philos. Psychol.* 1, 211–224. doi: 10.1007/s13164-009-0010-7

- Müller, D., Gerber, B., Hellstern, F., Hammer, M., and Menzel, R. (2000). Sensory preconditioning in honeybees. *J. Exp. Biol.* 203, 1351–1364.
- Noe, A. (2005). *Action in Perception*. Boston, MA: MIT Press.
- Northoff, G. (2013). *Unlocking the Brain*. Vol. 2. *Consciousness*. New York, NY: Oxford University Press.
- Oberauer, K. (2001). The explanatory gap is still there. *Behav. Brain Sci.* 24, 996–997. doi: 10.1017/S0140525X01480113
- O'Regan, J. K., and Noe, A. (2001). A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 939–973. doi: 10.1017/S0140525X01000115
- Pace-Schott, E. F., and Hobson, J. A. (2002). The neurobiology of sleep: genetics, cellular physiology and subcortical networks. *Nat. Rev. Neurosci.* 3, 591–605. doi: 10.1038/nrn895
- Pantke, K. H. (2017). *Locked-in: Trapped in One's Own Body*. Berlin: Christine Kühn Foundation.
- Pika, S., and Mitani, J. (2007). Referential gestural communication in wild chimpanzees (*Pan troglodytes*). *Curr. Biol.* 16, R191–R192. doi: 10.1016/j.cub.2006.02.037
- Pollick, A. S., and de Waal, F. (2007). Ape gestures and language evolution. *Proc. Nat. Acad. Sci. U.S.A.* 104, 8184–8189. doi: 10.1073/pnas.0702624104
- Popper, K. R. (1963). *Conjectures and Refutations*. London: Routledge and Kegan Paul.
- Requin, J., Somjen, A., and Bonnet, M. (1984). “Bernstein's purposeful brain,” in *Human Motor Actions - Bernstein Reassessed*, ed H. Whiting (Amsterdam: North Holland), 467–504.
- Revonsuo, A. (1995). Consciousness, dreams and virtual realities. *Philos. Psychol.* 8:35. doi: 10.1080/09515089508573144
- Riesenhuber, M., and Poggio, T. (1999). Are cortical models really bound by the “binding problem”? *Neuron* 24, 87–93. doi: 10.1016/S0896-6273(00)80824-7
- Rosen, R. (1985). *Anticipatory Systems: Philosophical, Mathematical and Methodological Foundations*. Oxford; New York, NY; Toronto, ON; Sydney, NSW; Paris; Frankfurt: Pergamon Press.
- Schrödinger, E. (1944). *What is life?* Cambridge: Cambridge University Press.
- Searle, J. R. (2000). Consciousness. *Annu. Rev. Neurosci.* 23, 557–578. doi: 10.1146/annurev.neuro.23.1.557
- Seidel, R. J. (1959). A review of sensory preconditioning. *Psychol. Bull.* 56, 58–73. doi: 10.1037/h0040776
- Seth, A. K., Baars, B. J., and Edelman, D. B. (2005). Criteria for consciousness in humans and other mammals. *Conscious. Cogn.* 14, 119–139. doi: 10.1016/j.concog.2004.08.006
- Shanahan, M. (2005). Global access, embodiment, and the conscious subject. *J. Conscious. Stud.* 12, 46–66.
- Shevrin, H., and Dickman, S. (1980). The psychological unconscious: a necessary assumption for all psychological theory? *Am. Psychol.* 35, 421–434. doi: 10.1037/0003-066X.35.5.421
- Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations. *Neuron* 24, 49–65. doi: 10.1016/S0896-6273(00)80821-1
- Stapleton, M. (2016). “Leaky levels and the case for proper embodiment,” in *Embodiment in Evolution and Culture*, eds G. Etzelmüller and C. Tewes (Tübingen: Mohr Siebeck), 17–30.
- Stickgold, R., Hobson, J. A., Fosse, R., and Fosse, M. (2009). Sleep, learning, and dreams: off-line memory reprocessing. *Science* 294, 1052–1057. doi: 10.1126/science.1063530
- Suddendorf, T. (2013). *The Gap*. New York, NY: Basic Books.
- Suddendorf, T., and Busby, J. (2003). Mental time travel in animals? *Trends Cogn. Sci.* 7, 391–396. doi: 10.1016/S1364-6613(03)00187-6
- Suddendorf, T., and Busby, J. (2005). Making decisions with the future in mind: developmental and comparative identification of mental time travel. *Learn. Motiv.* 36, 110–125. doi: 10.1016/j.lmot.2005.02.010
- Talk, A. C., Gandhi, C. C., and Matzel, L. D. (2002). Hippocampal function during behaviorally silent associative learning: dissociation of memory storage and expression. *Hippocampus* 12, 648–656. doi: 10.1002/hipo.10098
- Tallon-Baudry, C. (2004). Attention and awareness in synchrony. *Trends Cogn. Sci.* 8, 523–525. doi: 10.1016/j.tics.2004.10.008
- Taylor, A. H., Hunt, G. R., Holzhaider, J. C., and Gray, R. D. (2007). Spontaneous metatool use by New Caledonian crows. *Curr. Biol.* 17, 1504–1507. doi: 10.1016/j.cub.2007.07.057
- Tewes, C. (2016). “Embodied habitual memory formation: Enacted or extended?” in *Embodiment in Evolution and Culture*, eds G. Etzelmüller and C. Tewes (Tübingen: Mohr Siebeck), 31–56.
- Tolman, E. C., and Gleitman, H. (1949). Studies in learning and motivation. I. Equal reinforcement in both end-boxes, followed by shock in one end-box. *J. Exp. Psychol.* 39, 810–819. doi: 10.1037/h0062845
- Tolman, E. C., and Postman, L. (1954). Learning. *Annu. Rev. Psychol.* 5, 27–56. doi: 10.1146/annurev.ps.05.020154.000331
- Treisman, A. (1999). Solution to the binding problem: progress through controversy and convergence. *Neuron* 24, 105–110. doi: 10.1016/S0896-6273(00)80826-0
- Tschacher, W., and Bergomi, C. (eds.). (2011). *The Implications of Embodiment: Cognition and Communication*. Exeter: Imprint Academic Press.
- Tulving, E. (1985a). How many memory systems are there? *Am. Psychol.* 40, 385–398. doi: 10.1037/0003-066X.40.4.385
- Tulving, E. (1985b). Memory and consciousness. *Can. Psychol.* 26, 1–12. doi: 10.1037/h0080017
- Tulving, E. (1987). Multiple memory systems and consciousness. *Hum. Neurobiol.* 6, 67–80.
- Turvey, M. T. (1990). Coordination. *Am. Psychol.* 45, 938–953. doi: 10.1037/0003-066X.45.8.938
- Vanderwolf, C. H. (2000). Are neocortical gamma waves related to consciousness? *Brain Res.* 855, 217–224. doi: 10.1016/S0006-8993(99)02351-3
- van Duijn, M., and Bem, S. (2005). On the alleged illusion of conscious will. *Philos. Psychol.* 18, 699–714. doi: 10.1080/09515080500355210
- van Hemmen, J. L. (2014). Neuroscience from a mathematical perspective: key concepts, scales and scaling hypothesis. *Biol. Cybern.* 108, 701–712. doi: 10.1007/s00422-014-0609-3
- Varela, F. J., Thompson, E., and Rosch, E. (1991). *The Embodied Mind*. Cambridge, MA: MIT Press.
- Vonasch, A. J., and Baumeister, R. F. (2013). Implications of free will beliefs for basic theory and societal benefits: critique and implications for social psychology. *Br. J. Soc. Psychol.* 52, 219–227. doi: 10.1111/j.2044-8309.2012.02102.x
- von Holst, E., and Mittelstaedt, H. (1950). “The reafference principle. Interaction between the central nervous system and the periphery,” in *Selected Papers of Erich von Holst: The Behavioural Physiology of Animals and Man* (London: Methuen), 39–73.
- von Uexküll, J. (1970). *Streifzüge Durch die Umwelten von Tieren und Menschen*. Frankfurt am Main: Fischer Verlag.
- Vygotsky, L. S. (1978). “The role of play in development,” in *Mind in Society: The Development of Higher Psychological Processes*, eds M. Cole, V. John-Steiner, S. Scribner, and E. Souberman (Cambridge: Harvard University Press), 92–104.
- Watanabe, S., and Huber, L. (2006). Animal logics: decisions in the absence of human language. *Anim. Cogn.* 9, 235–245. doi: 10.1007/s10071-006-0043-6
- Wegner, D. (2002). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- Weir, A. A. S., Chappell, J., and Kacelnik, A. (2002). Shaping of hooks in New Caledonian crows. *Science* 297:981. doi: 10.1126/science.1073433
- Whitehead, A. N. (1911). *An Introduction to Mathematics*. New York, NY: Holt.
- Whorf, B. L. (1962). *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. New York, NY; London: The Technology Press of Massachusetts Institute of Technology.
- Wittgenstein, L. (1963). *Tractatus Logico-Philosophicus, 17th Edn.* Frankfurt: Suhrkamp.
- Wittgenstein, L. (1996). *Philosophische Grammatik*. Wien: Springer.
- Wittgenstein, L. (2001). *Philosophische Untersuchungen: Kritisch-genetische Edition*. Frankfurt am Main: Suhrkamp.
- Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. *Anim. Behav.* 63, 293–299. doi: 10.1006/anbe.2001.1914

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Kotchoubey. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.