Functionalist Counterargument Against Nonemergent Theories of Consciousness (Physicalism and Nonemergent Panprotopsychism)

Alexander Kiselnikov

ABSTRACT

This paper deals with criticism of nonemergent theories of mind from the perspective of functionalism. The most important part of the criticism is the analysis of the basic tenet of functionalism, presuming that informational processes in the brain are a changing pattern of neurons activation and deactivation. One of the main ideas of the functionalist counterargument is that in the physical world such characteristic as neurons activation or deactivation does not exist, it can only be subjectively "collected" of really existing physical attributes of particles that form a neuron. Which means an objectively changing neuronal pattern of activation and deactivation is only changing physical attributes of a multitude of particles, while neuron activation and deactivation is only being "invented" subjectively. But out of the nonexistence of a changing pattern of neurons activation and deactivation results nonexistence of informational processes what in the end within the confines of functionalism would mean illusive existence of consciousness itself. And only if the attribute of activation and deactivation becomes something emergent then informational processes, along with consciousness become something really existing.

Key Words: Hard Problem of Consciousness, Functionalism, Information, Constitutional Panprotopsychism, Emergent Panprotopsychism, Physicalism, Emergent Functionalism

DOI Number: 10.14704/nq.2018.16.11.1127 **NeuroQuantology 2018; 16(11):104-110**

Introduction

What are informational processes (in the brain) and do they really exist in the physical world? According to Shannon's definition information is the choice of one condition out of many possible (Shannon, 1948), that's why a neuron only by being in the physical condition of activation or deactivation makes automatically a choice between two possible conditions, which means it starts to carry one bit of information. And then informational processes in the brain are a changing neuron activation and deactivation pattern, that's why they seemingly exist in the physical world. But to respond to this question with confidence we must challenge both possible alternatives, so let's distinguish the subjective information theory (informational processes do not exist in the physical world) and objective information theory (they do exist).

Objective information theory (structural and emergent)

Objective information theory suggests that informational processes do exist in the physical world. But if so, in the physical world there should exist the pattern of neuron activation and deactivation. That could be possible if the neuron is an emergent physical object that can be in two emergent physical conditions of activation and deactivation. Let's call informational processes that are changing emergent neuron attributes *emergent*.

Corresponding author: Alexander Kiselnikov
Address: National Research University - Higher School of Economics (Faculty of Humanities, School of Philosophy), Russian Federation
e-mail ✉ oscolok1543@gmail.com
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Received: 12 September 2017; **Accepted:** 18 August 2018

But maybe the attribute of neuron activation and deactivation is not emergent, but at the same time does exist in the world. Because of specific disposition of physical particles in the brain, their physical properties can build the attribute of activation and deactivation. If we take a look at the brain, its physical particles are already placed in such a way, that they are "combined" in certain groups, and each of those groups due to the physical impact on it caused by another group of particles changes its physical properties and causes physical impact on other groups of identic particles structure. That's why these groups can be in two physical conditions. If one particle group is under physical impact, and then it goes over to another group, it means this group is in condition 1, and if it doesn't go over, then the group is in condition 2. Neuroscientists call these particle groups *neurons*, and these two conditions – *activation and deactivation*, but this isn't a subjective metaphor, because the physical world is in fact divided in groups of particles, that can be in two conditions. Physical properties of physical particles placed in a certain way (that subjectively for us form a neuron), do really carry the choice of one physical condition of two possible. Informational processes in the brain are changing physical properties of physical particles placed in a certain order and do not have to be something emergent. We shall call these nonemergent informational processes *structural* processes.

Subjective information theory (subjective informational processes)

Informational processes exist in physical world, if there exists the pattern of neuron activation and deactivation, but both, the neuron and his possible conditions are subjectively "created" by an external observer. Of all physical particles a neuroscientist distinguishes only a certain group that he calls *neuron* or *transistor*, but the same particles could be arranged in absolutely different groups that would form different physical objects. Which means in the physical world only particles exist that could be arranged in different groups in different ways, that's why neither the neuron nor its attribute of activation and deactivation can be found in the physical world. Informational processes that are changing pattern of neuron activation and deactivation exist only in an external observer's mind. One could object to the subjective information theory that physical processes in a computer and in the brain make meaningful calculations regardless of the existence of an external observer, but we could answer that both in the brain and in a computer run physical processes whose results human interpret in the way most efficient for them.

Compatible with the subjective information theory can only be the philosophical standpoint of qualia illusiveness, suggested by Dennett (not the whole consciousness theory by Dennett, but this idea only). Informational processes are being devised by an external observer, but in reality, they don't exist, but in the same way consciousness and qualia according to Dennett's theory are subjectively imagined, while they do not exist in reality (Dennett, 1991). Informational processes and consciousness are two illusions not by coincidence, they are one and the same illusion, while in reality only changing physical attributes of brain particles exist. That's why to discard completely the subjective information theory we will have to reject the idea of illusiveness of qualia and consciousness as well.

As a result, we have to answer the following question: are the informational processes that run in the brain subjective, structural or emergent?

Functionalist counterargument against nonemergent consciousness theories

Functionalist counterargument will rest on two premises, from which emergence of the activation and deactivation attribute will result:

information in the brain is being coded by using the neuron activation and deactivation attribute (the neuron can be in two possible physical conditions and its constant being in one of those conditions automatically is a choice of one condition of two possible) (this premise doesn't include the concept of the obligatory neuron emergence and the activation and deactivation attribute; the neuron can be only a special configuration of physical particles).

consciousness is information (not necessarily information only, but information in particular)

These two premises are the base of functionalism, that's why I called the counterargument against nonemergent consciousness theories *functionalist*. At the end of the proving, starting out from these premises I will come to the conclusion that functionalism automatically means emergentism, what almost all functionalists will strongly disagree with; and still I insist that it results from the adduced premises. That's why either the premises are wrong or the logical transitions in the counterargument are

wrong, or functionalism is after all emergent functionalism as a matter of fact.

Consciousness is information, but despite Shennon's definition of information, it is still not completely clear "what it really is". The brain consists of 86 billion neurons, every neuron can be activated or deactivated, that's why the brain information space can be imagined as a space that includes 86 billion subspaces, in every single of them happens a choice of on variant of two possible. But the choice of one "what"? One real physical condition of two possible or the neuron pattern refers to the choice of one non-substantional condition of two possible?

In such a big information space it is very hard to trace relationship of information and consciousness, so first we shall examine this issue in a sketchy variant of brain that will consist of four neurons only, then on this simplified example their interconnection will become most obvious and we shall return to the real brain. For the convenience of reasoning we will also distinguish two levels on which neuron can be examined: micro and macro. The micro-level is common physical neuron description in terms of particles and their physical attributes, while on macro-level the neuron is in the condition of activation or deactivation.

If each of the four neurons on macro-level is in one of two conditions of activation or deactivation, then the whole little 4-neuron-brain on macro-level can be in 16 different conditions (2^4). Which means that every neuron when in activation or deactivation condition was a choice of one condition of two possible and the pattern of activation and deactivation of all four neurons will also be the choice of one condition of 16 possible. As well as each of the four neurons in its activating-deactivating structure carries the opposite unrealized opportunity (if the neuron is activated then it is not deactivated and the other way round). This is why the pattern of four neurons in its activating-deactivating structure will differ from 15 other possible non-realized patterns. Because of the fact that neurons become information carriers, they additionally start to carry this difference in themselves.

But now let's focus on the fact that the neuron pattern carries the choice of a certain physical substantional condition out of 16 different substantional conditions, but we can imaginatively "move away" from here the physical substantionality and it appears that it just carries the choice of one "non-substantional" of 16 different "non-substantionals". Let's clear this point by the following example: let's imagine that do exist 16 different colors, 16 different physical conditions of electron and 16 different animal species. And despite the fact that these three groups of elements do differ from each other, it is possible to find an abstract "formula", that would include the everything they have in common. This "formula" will be 16 different "unfilled places for differences". It can become 16 different colors, electron physical conditions or animals, but for now it is only 16 different empty cells. And the neuron pattern, being a choice of one physical condition of 16, simultaneously carries the choice of this one "non-substantional" out of 16 different non-substantionals. While finally there is no clear understanding of what this "non-substantional" is, but clear, separated from the information carrier is exactly such a "non-substantional choice" of one "empty", that differs from 15 other "empty"

That is why the idea of consciousness being information can mean only two possible alternatives: either it is a choice of a certain physical substantional neuron pattern of 16 possible, which means it is a neuron pattern, carrying the difference from 15 other possible patterns, or consciousness is only the choice of one non-substantional condition, different from other 15 non-substantional (the brain that consists of 86 billion neurons can be in 2^{86bln} possible conditions; that's why real consciousness will be the choice of one different from other (2^{86bln}-1).

Let's examine both possible variants:

1. *consciousness is information only (the choice of a non-substantional condition)*

As we have already mentioned, information is a special "formula" for choosing one of 16, a formula we can plug anything in (the choice of one of 16 colors of one of 16 emotions), but it is also possible that we do not "plug in" anything, and then it remains a choice of an "empty cell" of 16 different "empty cells". Which means consciousness at every moment of time is an element in a special mathematical set that consists of 16 different "non-substantional empty elements" (let's call this set *phenomenal*). But if consciousness is merely one non-substantional condition, different from other 15, then arises the question: what is this "non-substantional condition"? And how does this non-substantional condition differ from "nothing", "emptiness"? And this question must be answered: the only difference is that it is "nothing" that carries the difference from 15 other "nothings". The usage of the term "nothing" may seem

nonacceptable, but European philosophy has been doing it for a long time already (Hegel, 1997; Sartre, 1984). That's why we should either completely give up using the idea of information as differences, without things that are different, or say that differences without things that are different are differences between "nothings".

But if consciousness is pure information only, then we cannot say that it exists (because it is just «nothing» different from other "nothings"), on the other side, to the extent that it exists it is a pattern of neuron activation and deactivation (that carries the choice of one of these "nothings"). That's why if consciousness were information only, then really in some sense Dennett would be right when arguing that it is illusion.

But consciousness isn't "nothing", containing only the characteristic of being different from other 15 (2^{86bln}-1) "nothings". Condition of consciousness, when it sees the qualia of red is not only not some "empty, non-substantional element 1", not some "empty, non-substantional element 2" and ... not some non-substantional 15^{th}, but at least not "yellow ","not orange" and ..."not purple". That is, the state of consciousness, when it contains the qualia of red, is not only an empty, non-substantional element, different from other empty non-substantional elements, not only "nothing" other than the "nothing", but also the substantial chromaticity, redness. And to this it can also be added that the condition of consciousness, when it sees the qualia of red, is "nothing", different from the other "nothing" and its condition, when it sees the qualia of green, it is also only "nothing", different from other "nothings", so if there were no substantiality in qualia (that is, it would be an illusion), then it would be incomprehensible "why the illusory qualia of red and green subjectively appear to be different from each other".

Therefore, the substantionality must overlap the phenomenal set and turn it into 16 (2^{86bln}) different phenomenal conditions, in particular including different colors. Consciousness is not a pure information, but information, which is filled with some substantionality. This substantionality can be physical (physicalism), protophenomenal (not emergent panprotopsychism) or phenomenal (dualism, emergentism). If we prove that physicalism and not emergent panprotopsychism are not true, then only the last variant will remain.

2. consciousness is information, filled by substantionality

In the beginning, we'll consider the option, when consciousness is information, which is filled with physical substantionality (let's call it functionalistic physicalism). The neuron is not an emergent physical object, but the brain consists of certain "groups of physical particles" that are always in the same physical condition of two possible ones. Therefore, the physical condition of 86 billion of such groups of physical particles is the choice of one of this physical condition of 2^{86bln} possible. And then consciousness is a changing physical property of 86 billion particle groups, different from their other (2^{86bln}-1) unrealized physical conditions. Qualia is only a physical property that carries the difference from its other possible physical conditions, and nothing else.

At the same time, information can be filled not only with physical, but also with protophenomenal substantionality (let's call it functionalist panprotopsychism). Russell suggested that material particles from the outside look like something physical, but from the inside "objectively" they are something psychic, that is, their physical properties are in reality some phenomenal properties (Russel, 1954). Chalmers suggested calling these properties protophenomenal to emphasize their difference from a more complex phenomenal experience (Chalmers, 1996). Depending on the relationship of these protophenomenal and phenomenal properties panprotopsychism is divided into emergent and not emergent.

For a functionalist non-emergent panprotopsychism consciousness is information that is "filled" with protophenomenal substantionality. The brain consists of protophenomenal particles, a special configuration of which unites them into 86 billion groups, which all the time are in one protophenomenal condition of two possible. Therefore, the changing protophenomenal properties of brain particles begin to carry informational processes, which include the continuous selection of one protophenomenal condition from the set of possible ones. And in this case consciousness is the changing protophenomenal properties of brain particles, different from their other (2^{86bln}-1) nonrealized protophenomenal conditions. Qualia is only a protophenomenal property, carrying in itself the difference from its other possible protophenomenal conditions, and nothing more.

At the same time, David Chalmers, in the framework of panprotopsychism, puts forward another theory of consciousness (constitutive panprotopsychism), which he also considers nonemergent. Chalmers suggests that a special

combination of protophenomenal properties can "constitute" phenomenal properties in such a way that these phenomenal properties will not be protophenomenal themselves (Chalmers, 2017). That is, the protophenomenal properties are not phenomenal, but their special combination "without an emergent transition" becomes phenomenal. In other words, behind the word "constitute" emergent transition "is hiding", which means, constitutive panprotopsychism is not an non-emergent theory of consciousness. Therefore, the only functionalist non-emergent theories of consciousness are physicalism and non-emergent panprotopsychism (presented exactly as they were formulated in this section).

"hard problem of consciousness" in the language of the theory of sets

And now, before moving to the final stage of the functionalist counterargument, it is necessary to answer the following question: Is it possible to imagine the physical world (more specifically, the brain) without substantionality? What will it be? For example, if the physical world consisted of only one electron that could be in two physical states, then what would this electron be without substantionality? If we remove substantionality, then there will be neither the electron nor two of its physical states, but there will be only two possible non-substantional conditions, two "nothing" that are different from each other. The physical world, consisting of two electrons in two conditions, will be nothing, different from the other three nothing. Therefore, the physical world without substantionality will also be "nothing", different from the other "nothing".

Now let's return to the simplified version of the brain, which consists of only four neurons and will represent it without substantionality. Each particle that forms a neuron can have different values of its physical parameters, so the entire neuron can be in a huge number of different states (let's say "n"). Then a system of four neurons can potentially be in n^4 different physical conditions. Therefore, the brain without substantionality will be an empty, non-substantional element, "nothing" that carries a difference from other (n^4-1) empty elements. A system of n^4 mutual "empty elements" is called a "physical set". If the physical substantiality is "superimposed" on this "nothing", different from other (n^4-1) "nothing", then it will be filled with physical qualities and will "turn" into a physical world. And it does not even matter whether the physical world without substantionality is pure information or not, because the idea of the physical world as pure information on which physical substance is superimposed will carry in itself all the properties of the physical world (which is what is needed for further proof) .

Now there are two mathematical sets:

The set W is a physical set consisting of n^4 different empty elements.

The set F is a phenomenal set consisting of 2^4 different empty elements.

The physical brain without substantionality is the choice of one element from the physical set, and consciousness without substantionality is the choice of one element from the phenomenal set. Therefore, if a physical set includes a phenomenal set, then the choice of one element from the physical set will include the selection of one element from the phenomenal set. And then physical (or protophenomenal) substantionality, superimposed on the physical set, will fill with physical (or protophenomenal) substantionality not only the physical brain, but also consciousness. That is, if in the mental "removal" of the substantionality of the physical brain, its remaining "frame" will carry in itself a phenomenal set with the choice of one element from it, then the "reverse" filling with the substantionality of this "frame" of the physical world will automatically lead to the fact that consciousness will be filled with substantionality as well. In the end, if the physical set includes the phenomenal, then consciousness is a physical (or protophenomenal) property.

The difficulty in answering this question lies in the special structure of the physical set W. On the one hand, it corresponds to the n^4 physical states of the four-neural brain, but at the same time at the macro level this brain is in 16 different states of activation and deactivation. Because of this, all the physical states of the small brain can be divided into 16 subgroups in such a way that within the subgroup different physical states from each other will "realize" the general macro pattern of activation and deactivation. Therefore, the n4 elements of the physical set B can be divided into 16 subsets, but to do so the characteristic based on which the selection of the subset will occur should exist in the physical set itself. This sign is a pattern of activation and deactivation of the brain, which is formed by separate properties of activation or deactivation. According to the structural theory of information, this property of activation and deactivation is formed by a special configuration of

physical properties of brain particles, so it can be in the physical world itself (and hence in the set W). But is the structural theory of information true?

In order to answer this question, it is necessary to pay attention to the nerve cell. If a neuron receives a strong electrical signal from another neuron, the charge of its membrane rises above -70mV and sharply "collapses" to 0. This will lead to the fact that an electric signal will go through the membrane to the axon, which, reaching the synapse, will cause releasing mediators and transmitting a signal to the neighboring neuron. But how exactly do we understand which physical state of the neuron corresponds to its macro-condition of activating, and which of deactivating? We can say that the physical charge of the neuron membrane below -70 mV is associated with deactivation, and about 0 and V with activation. But you can also say that if the electrical signal that came to the neuron passed on to another neuron, then the physical state of this neuron is associated with activation, and if not, then with deactivation. That is, there is a certain criterion, and if the current physical state of the neuron corresponds to it, then it belongs to group 1 (activation), and if not, to group 2 (deactivation). And these two different criteria for separating the states of a neuron into two groups seem to carry the same meaning, but in reality there will be cases when they will refer the same state of a neuron to different groups. And even the strict second criterion, which seems to be unable to be mistaken, can be questioned. For example, if we disable the axon for neuron 1, but when the membrane charge approaches the disconnected axon, we still transmit the signal to the neighboring neuron 2 by stimulating it with an electrode, will in this case neuron 1 be considered activated or not? Or when a neuron 1 signal comes to the neuron 2, it is possible to block the signal even before it activates the neuron 2 and activate the neuron 3 through an electrode the axon of the neuron 2 is reaching out for. Is then the neuron 2 activated or deactivated? And if we do this to all the neurons in the brain?

And even if it turns out to find an absolutely ideal criterion for assigning a neuron's physical state to an activation or deactivation group, then one can even single out such a strange criterion as: if the membrane charge is a multiple of 2, then we refer the state of the neuron to the 1st group (activation), and if no, then to the second (deactivation). The names of the two groups of states "activation" or "deactivation" are mere conditionality, and it is important only that, being in one of the two groups, the physical state of the neuron carries the choice of one of this group of two possible, that is, it carries a bit of information. Therefore, any two groups of physical states of the neuron at the information level will have the same meaning as the two activation and deactivation groups. That is, in reality there is no activation or deactivation at all, and there are only two possible states of the neuron. Therefore, the divisibility or non-divisibility by 2 of the charge of the membrane can also be a criterion for dividing all possible neuron conditions into two groups. And then within the framework of the structural theory of information, the divisibility or non-divisibility by 2 will also be connected with some special configuration of physical particles forming a neuron, and it can also be said that it exists in the physical world itself. Or if the physical charge of a neuron's membrane is divisible or non-divisible by 3, 4, 5, and so on, then the physical states of the neuron can also be divided into two groups of states. Such "meaningless" criteria can be invented infinitely and eventually it will turn out that the physical states of a neuron can be divided into two groups in an infinite number of ways. And then always being in one physical state of two possible, the neuron will be the choice of one of this state from two possible ones, that is, it will carry a bit of information in itself. And this means that the same neuron by criterion 1 will be in the activation state, by criterion 2-in deactivation, by criterion 3 - again in activation and so on ad infinitum. As a result, in the same brain, on the same neurons, there will be an infinite number of different information processes. For us, they will be meaningless and chaotic, but for the physical world they will be as possible as those information processes that neuroscientists study. In addition, you can even come up with a physical system, in which information processes will produce meaningful calculations immediately by two criteria of dividing the carriers of its information into two groups of possible physical states. Therefore, we can conclude that the criterion based on which the states of a physical system (neuron) are divided into two groups does not exist in the physical world itself, but in our consciousness only.

But if in the physical world there is no single division of the physical states of the neuron into two definite groups, then it will not exist in the physical set W. In this case, the set W reflects only the physical properties of the particles forming the neurons, and therefore, it does not contain its

"breaking" into 16 subsets. And then the structure of the set W is simply all the elements' difference from each other. Each element in the set F has the property to be different from 15 other elements, and the element in the set W has the property to be different from (n^4-1) other elements. The sets W and F consist of elements that have different properties, and therefore they consist of different elements, which means that the physical set W can not include the phenomenal set F.

In the end, if a physical (protophenomenal) substance is superimposed on a physical set, then a physical (protophenomenal) brain appears. But since the physical set W did not include the phenomenal set F, consciousness did not appear in the physical (protophenomenal) brain. Physical particles appeared, whose states can be further broken down into two groups of activation and deactivation, thereby creating a phenomenal set of F, but in the physical world this division does not exist. The physical set does not include the phenomenal, which means that consciousness can not exist in the physical (protophenomenal) world. And this local proof for a set of W and F consisting of 2^4 and n^4 elements is easily transfered to the real brain, where W and F consist of 2^{86bln} and n^{86bln} elements, and everything happens there in exactly the same way. Physicalism and non-emergent panprotopsychism can not be true.

Conclusions

In conclusion, if a neuron becomes an emergent object that is constantly in one of two emergent states of activation or deactivation, then this emergent property of a neuron makes real information and simultaneously "fills" it with an "emergent" substantionality. This is possible in the context of emergent materialism or emergent panprotopsychism, but the second option looks more articulate. The protophenomenal properties of the neuron particles form a new emergent property that from the outside looks like neuron activation or deactivation, and from the inside "objectively" is one of two phenomenal states. A living cell for biology is a unit of life, so there is nothing strange about it if the unit of the phenomenal becomes a nerve cell. In addition, all living cells to some extent receive through the membrane signals from the outside world and can somehow answer them or not, so, probably, all living cells are in two emergent states of activation (response to external stimulation) or deactivation (the absence of this answer). That is, I believe that all living cells do have phenomenal.

But of all living cells, neurons are the most "alive" ones. Neurons not only have dendrites that allow them to receive signals from thousands of other cells, but they also maintain themselves in a state of maximum instability and disequilibrium, which makes them sensitive even to very weak external signals. This is what allows neurons to begin to specialize in maximally deep "joint" analysis of external stimulation and constructing responses to it. When neurons are combined into a system, they have a unite emergent pattern of activation and deactivation, which "from within" is that phenomenal state that through billions of years of evolution will become a phenomenal consciousness.

References

Chalmers DJ. The conscious mind: In search of a fundamental theory. New York: Oxford University Press, 1996.

Chalmers DJ. The Combination Problem for Panpsychism. In (G. Bruntrup and L. Jaskolla, eds.) Panpsychism: Contemporary Perspectives. Oxford: Oxford University Press, 2017.

Dennett DC. Consciousness Explained. Boston: Little, Brown, 1991.

Hegel GW. Hegel's Science of Logic, trans. AV Miller. Amherst: Humanity Books, 1969; 99: 37.

Russell B. The Analysis of Matter. New York: Dover, 1954.

Sartre JP. Being and Nothingness. trans. Hazel E. Barnes. New York: Washington Square Press, 1984.

Shannon CE. A mathematical theory of communication. Bell Systems Technical Journal 1948; 27: 379-423.