Metadata of the chapter that will be visualized in SpringerLink

Book Title	Philosophy and Theory of Artificial Intelligence 2017
Series Title	
Chapter Title	Agency, Qualia and Life: Connecting Mind and Body Biologically
Copyright Year	2018
Copyright HolderName	Springer Nature Switzerland AG

Corresponding Author	Family Name	**Longinotti**
	Particle	
	Given Name	**David**
	Prefix	
	Suffix	
	Role	
	Division	
	Organization	
	Address	Columbia, MD, USA
	Email	longinotti@hotmail.com

Abstract

Many believe that a suitably programmed computer could act for its own goals and experience feelings. I challenge this view and argue that agency, mental causation and qualia are all founded in the unique, homeostatic nature of living matter. The theory was formulated for coherence with the concept of an agent, neuroscientific data and laws of physics. By this method, I infer that a successful action is homeostatic for its agent and can be caused by a feeling - which does not motivate as a force, but as a control signal. From brain research and the locality principle of physics, I surmise that qualia are a fundamental, biological form of energy generated in specialized neurons. Subjectivity is explained as thermodynamically necessary on the supposition that, by converting action potentials to feelings, the neural cells avert damage from the electrochemical pulses. In exchange for this entropic benefit, phenomenal energy is spent as and where it is produced - which precludes the objective observation of qualia.

Agency, Qualia and Life: Connecting Mind and Body Biologically

David Longinotti(✉)

Columbia, MD, USA
longinotti@hotmail.com

Abstract. Many believe that a suitably programmed computer could act for its own goals and experience feelings. I challenge this view and argue that agency, mental causation and qualia are all founded in the unique, homeostatic nature of living matter. The theory was formulated for coherence with the concept of an agent, neuroscientific data and laws of physics. By this method, I infer that a successful action is homeostatic for its agent and can be caused by a feeling - which does not motivate as a force, but as a control signal. From brain research and the locality principle of physics, I surmise that qualia are a fundamental, biological form of energy generated in specialized neurons. Subjectivity is explained as thermodynamically necessary on the supposition that, by converting action potentials to feelings, the neural cells avert damage from the electrochemical pulses. In exchange for this entropic benefit, phenomenal energy is spent as and where it is produced - which precludes the objective observation of qualia.

1 Introduction

The thesis of strong artificial intelligence is that the mind is essentially a computer, such that a suitably designed and programmed machine could pursue its own goals and have phenomenal experiences (Johnson-Laird 1988). In this paper, I contend that these claims are analytically and scientifically untenable, and describe a biological solution to the mind body problem. My approach is naturalistic and scientific; I assume that agency and qualia supervene on other phenomena that we take to be natural, and that qualia have regular, discoverable effects on the world. The theory I offer is based on the evaluation of hypotheses for their coherence with the concept of an agent, empirical data and laws of physics. Scientific explanation often requires the postulation of mechanisms, like the events by which an axon conducts an electro-chemical pulse (Machamer et al. 2000). Accordingly, the consideration of mechanisms is central to my method, which leads me to infer that actions and feelings have a common origin in the homeostatic nature of living matter.

The three main sections of the paper concern life, agency and qualia, respectively. I first review the relevant properties of a living system as an entity that is self-organized,

D. Longinotti—Independent.

V. C. Müller (Ed.): PT-AI 2017, SAPERE 44, pp. 1–14, 2018.
https://doi.org/10.1007/978-3-319-96448-5_6

and that maintains itself against thermodynamic decay. The next section concerns the nature and source of agency. From the concept of an agent, I deduce that it is a living substance. Because behavior motivated by a feeling has the homeostatic form of an action, I infer that life is the source of qualia, and that mental causation is based in the regulatory function of affective experiences. In the third main section I address the nature of qualia and the mechanism of their production. Laws of physics are adduced for the hypothesis that a feeling depends on the matter and energy at its location, rather than a causal pattern. From empirical evidence, I surmise that qualia are a distinct form of energy, a property generated in specialized neurons. The subjectivity of qualia is explained as required by thermodynamics if, by producing them, the source of the qualia avoids an increase in its entropy. I conclude with some remarks on the merits of the theory.

I avoid the term "consciousness" in the paper due to its many meanings, one of which involves cognitive attention. Here, my theorizing on consciousness is limited to 'qualia', what Block (1995) describes as 'phenomenal consciousness.' I use "subjective" to mean that a quale is not objectively observable and, in that sense, is private to its subject.

2 Life

2.1 Life Is Self-organizing

The scientific view of life is that it is a natural phenomenon. A living cell is commonly characterized as self-organized, that is, the structure of the cell results from the materials that comprise it, not from an externally imposed design plan. Living matter is similar in this way to other substances that depend on chemical forces for their composition (e.g., crystals, acids, proteins). No outside influence is needed for the internal organization of such substances. As Pross (2003) explains, "living systems are no more than a manifestation of a set of complex chemical reactions and, as such, are governed by the rules of kinetics and thermodynamics." The relevant implication with regard to agency is that the behavior of a living organism in a particular environment is self-determined; its movements result from the way its constituent materials organized themselves.

2.2 Life Is Self-maintaining

Jonas (2001) writes that "in living things, nature springs an ontological surprise in which the world-accident of terrestrial conditions brings to light an entirely new possibility of being: systems of matter that are unities of a manifold … in virtue of themselves, for the sake of themselves, and continually sustained by themselves." Like all systems, a living organism obeys the second law of thermodynamics, which states that the entropy (i.e., disorder) of an isolated system increases with time. That is, every system tends to decay to its equilibrium state of maximum disorganization; for a living thing, this deterioration results in its death. Preventing or slowing this breakdown requires the expenditure of energy from outside the system. In this regard, a living cell functions somewhat like a refrigerator; it consumes energy from external sources to prevent thermal decomposition. However, a refrigerator only slows the decay of things inside it, while a cell sustains its own substance. Schrödinger (1944) views this capability as unique to living matter, and explains that "the device by which an organism maintains itself stationary at a fairly high level of orderliness (=fairly low level of entropy) really consists in continually sucking orderliness from its environment."

The life-supporting order that is obtained from the environment is 'free energy' in various forms, energy at a sufficiently low level of entropy such that it can be metabolized by the organism. For life on earth, the ultimate source of free energy is sunlight, which is used by plants to construct organic complexes that contain chemical energy. Much of the energy and material consumed by a living cell is used in re-synthesizing the numerous proteins required to maintain the cell, as the proteins continually degrade (Pross 2012). Systems that consume energy to maintain themselves in a far-from-equilibrium state are described as 'dissipative' by Prigogine (1978) in that they reduce the amount of free energy in the environment, the energy that can be used for work. Schneider and Kay (1994) hold that "life should be viewed as the most sophisticated (until now) end in the continuum of development of natural dissipative structures, from physical to chemical to autocatalytic to living systems."

Maturana and Varela (1980) characterize a living system as a mechanism that is homeostatic with regard to its own composition. They use the term "autopoietic" (i.e., self-constructing) for such a system: "an autopoietic machine continuously generates and specifies its own organization through its operation as a system of production of its own components … it has its own organization (defining network of relations) as the fundamental variable which it maintains constant."

So, a living cell is self-organized, and its movements are self-determined relative to its environment. Those movements involve the consumption of materials and energy to repair the structure of the cell against the effects of heat and other threats to its biological integrity. A living cell is a homeostatic (i.e., self-maintaining) substance.

3 Agency and Mental Causation

Conceptually, an agent is something that moves itself to realize a goal; such behavior is termed an action. The lack of the goal is the motivation for an action, and the movement for the objective is initiated and controlled by the agent itself. A successful action concludes with the attainment of the goal, which ends the motivation for the behavior.

3.1 An Agent Is a Type of Substance

An agent 'moves itself' in the sense that it determines the way it behaves in response to some stimulus. An agent is 'active'; its movement is powered by energy it contains. In the words of Barandarian et al. (2009), "an agent is a source of activity, not merely a passive sufferer of the effects of external forces."

In general, the two determinants of a system's movement are the characteristics (material and form) of its components, and their organization. Computers and the operations they perform are multiply realizable: the same sequence of computational operations (i.e., the algorithm or software program) can be implemented using a wide variety of materials, and the same material can be used to realize a limitless variety of computational algorithms.

Because a computer is multiply realizable, the specific sequence of operations it performs depends only on its organizational structure. But this structure is not determined by the material of the computer. If it were, the same type of material could not be used to run many different programs. So, the material composition and functional organization that determines how a computer moves is not intrinsic to the computer. It is not an agent, but a tool of its designer.

Accordingly, a necessary property of an agent is that it is self-organized, which makes it the source of its own behavior. This entails that the structure of an agent depends on forces that are intrinsic to its components. Hence, an agent is organized by chemical bonding, forces that inhere in the very nature of the joined materials. But when entities combine chemically, the resulting substance differs in kind from its constituents taken individually. For example, the characteristics of hydrogen and oxygen are lost when they bond to form water.

So, the concept of an agent entails that it is a chemically composed substance, one which consumes energy to move for a goal.

3.2 Agency Depends on Living Matter

What kind of chemical substance is an agent? An action commences with some sort of change *within* its agent, a change that disturbs the agent from its quiescent state. This change 'motivates' (i.e., is the proximate cause of) the action. But because an agent moves for a goal, it must also be the 'want' of the goal that triggers its movement. So, the want of the goal is the motivating change in the agent. Accordingly, the goal of an action is to undo the change in the agent that motivated the movement, thereby returning the agent to its prior 'resting' state. Hence, a successful action has a 'circular' form; it begins and ends in the same entity within the agent. In contrast, a reflex is a 'linear', programmed movement that, once initiated, is carried out irrespective of its effect (if any) on that which triggered it. Unlike a reflex, an action has a homeostatic nature; an agent moves to keep itself in a certain state. And, as argued above, it is a substance that determines its own movement. Hence, an agent is a material having a homeostatic nature.

The concept of an agent accords with the unique character of living matter. A deviation from its self-maintaining activity causes a living cell to expend energy such that, if its movement is effective, the cell returns itself to a more sustainable, dynamic state. I believe that Aristotle recognizes the homeostatic basis of agency where, in Apostle's (1981) translation of *de Anima,* he asserts that "the principle of moving and stopping … is a power of such a nature as to preserve that which has it and to preserve it qua such." Aristotle coins a word for this power: *entelecheia*. In his literal translation, Sachs (2001) takes this term to mean "being at work staying itself". This description of an agent is fully consistent with the scientific characterization of life as reviewed above, wherein a cell is depicted as consuming energy in a manner that maintains its material composition and structure - thereby enabling it to continue this very activity. A living cell is its own goal.

But living *organisms* do not necessarily behave in this way; a moth that flies into a flame apparently moves reflexively, rather than for self-preservation. How, then, do some living organisms move as agents?

3.3 Qualia Originate in Living Matter

A movement of an organism that is motivated by an affective feeling has the homeostatic form of an action; successful behavior ends the painful or pleasurable feeling. The usual response to thirst is an example; the feeling that motivates the movement is extinguished in the organism when it restores itself to its hydrated condition. Similarly, behavior for pleasure ends when satiation is reached. Damasio (2012) remarks that "in brains capable of representing internal states … the parameters associated with a homeostatic range correspond, at conscious levels of processing, to the experiences of pain and pleasure." On the assumption that a feeling is caused by some change in its subject, hedonically motivated movement is an action; attainment of the goal occurs when the part of the organism that produced the feeling is returned to its prior, resting state. In this regard, Spencer (1855) notes that feeling-related movements begin when reflexive motion ends: "…as the psychical changes become too complicated to be perfectly automatic, they become incipiently sensational. Memory, Reason, and Feeling take their rise at the same time." With my earlier inference that an action is a movement of living matter, the observation that hedonic feelings can motivate actions enables a straightforward deduction regarding the origin of at least some types of qualia:

Every action is caused by a change in a living substance.
Some actions are caused by affective feelings.
An affective feeling is caused by a change in a living substance.

This deduction is specific to hedonic feelings. But all qualia are subjective, and I will argue in Sect. 4.4 that subjectivity results from the living nature of the source of qualia. Assuming that is correct, it entails that all qualia - not just the affective types - depend on life. The syllogism above also presumes that qualia can influence behavior in some way. The question of how that occurs is the problem of mental causation.

3.4 Qualia Affect Behavior as Control Signals

For some, the claim that feelings can influence physical movement is equivalent to Cartesian interactionist dualism. This is the view that mind and body are fundamentally different, but that there are causal connections between them. Dualism is not entailed by interactionism, however. In Newton's time, many held that his theory of gravity required the existence of a supernatural phenomenon, because it was widely believed that all forces operated by contact (Gibbon 2002). The current, 'physicalist' view of the world reflects a stance similar to that of Newton's critics; physicalists typically claim that the 'physical' (i.e., non-mental) world is causally closed. But this is contrary to experience. If a phenomenon had no causal relationships with the rest of the world, we would be totally oblivious of it - but we are not oblivious to qualia. In Russell's (1959) view, they are the only sort of thing we know by direct 'acquaintance', rather than through inference.

On the theory of qualia offered here, they can have effects at two levels. At the micro-level of their production, qualia benefit the biological integrity of their living source, as I will posit in accounting for their subjectivity. At a higher level of organization, a phenomenal experience can prompt an organism to act in some way, as assumed in the previous section with the example of thirst.

A possible objection to the view that qualia can cause actions is that, if they were to influence an organism's movement, they would have to do so by exerting a telekinetic force on neural activity - and there is no evidence of such a force. But telekinesis is not necessary for feelings to affect behavior; they can do so as control signals. Analogously, a ship can be steered automatically using light from stars, even though the starlight exerts no relevant force on the ship. All that is required is that the ship be able to *detect* the stars, measure their positions relative to its heading, and adjust its course accordingly. All the force needed to change the direction of the ship is supplied by the ship itself, not by the stars. Similarly, no force on neural activity is needed for a feeling to affect the behaviour of an organism; the organism need only detect the feeling and respond to it in some way – generally, by selecting a type of movement that will influence the feeling (e.g., by eliminating the organism's thirst).

One source of the perceived difficulty in understanding mental causation is a line of reasoning that Kim (2005) calls the "supervenience argument". Let M be the experiential property of a mental state like pain, where M supervenes on its physical base P. M is thought to cause P*, some neural event that results in pain-reducing behavior. But P also appears to be the cause of P*, in which case P* is causally over-determined. Such dual causation is very unlikely so either M is reducible in some way to P, or M is epiphenomenal.

This argument posits that the neural state P, on which M supervenes, is also the cause of P*, the physical response to M. But this is generally not the case. Between the feeling and the behavioral response to it, there can be a lengthy interval of practical reasoning concerning the type of movement (if any) to perform. Otherwise, every movement would be a reflex. Hence, M supervenes on P, but P does not cause P*. M and P* have different causal bases, so causal over-determination is not entailed by M's supervenience on P. This can be seen with the ship analogy wherein one mechanism (a photo-detector) produces a control signal from the starlight, and a separate, mechanical system uses that signal to adjust the ship's rudder.

Hence, the science of mental causation is that of control theory (i.e., cybernetics), wherein the operation of a system is typically adjusted based on an error signal that represents the difference between the goal for the system and its actual state (Ashby 1956). A number of theorists have characterized goal-oriented behaviour as a process involving feedback control (MacKay 1966; Powers 1973; Carver 1979; Carver and Scheier 1981; Marken 2002). The 'navigation' of an organism using its feelings as control signals is similar to the stellar navigation of a ship – except that, in the case of the organism, the source of the feedback signals is internal to the 'vessel'. The organism experiences affective qualia and, using learned behavior and/or practical reasoning, responds accordingly. Just as the imagined ship can't navigate without the starlight, an organism that is guided by its feelings is in mortal danger without them. Humans that lack sensitivity to pain often die before reaching adulthood, because they fail to notice injuries (Nagasako et al. 2003).

The view that affective qualia perform a control function is not new to psychology; Cannon (1932) describes the role that feelings like hunger and thirst perform in the homeostatic regulation of bodily requirements - like water, sugar, proteins, fat and calcium, as well as the oxygen and salt contents of the blood. Schulze and Mariano (2004) offer the following, generalized account:

> Since hedonic states arise whenever a control system produces a chronic regulation error, this implies that the control system is unable to regulate an important physiological variable within the limits required to maintain the integrity of the organism. The hedonic states that arise in response to an increasing regulation error serve to co-opt the behavioral system and its resources. It is then up to the latter to select and execute the appropriate behaviors drawing on cognitive systems in the process.

In addition to physiological conditions, thoughts can also result in motivating feelings, as Hume (1739) describes:

> ’Tis obvious, that when we have the prospect of pain or pleasure from any object, we feel a consequent emotion of aversion or propensity, and are carry’d to avoid or embrace what will give us this uneasiness or satisfaction. ’Tis also obvious, that this emotion rests not here, but making us cast our view on every side, comprehends whatever objects are connected with its original one by the relation of cause and effect.

So, an action may be stimulated by the *anticipation* of pleasure or pain, and this apparently occurs through a faint experience of the expected feeling. Freud famously contends that this sort of process can occur subconsciously, causing us to pursue or repress particular thoughts and memories. Hence, affective feelings function as control signals that motivate an organism to think and/or move to realize a goal-state.

The ability to respond to their feelings conferred a significant biological advantage on those species that evolved this capability. An organism that is limited to reflexive movements is constrained by its evolutionary past, like the aforementioned moth that flies into a flame. In contrast, motivation by its affective feelings enables an individual organism to respond in the present, to new threats and opportunities. Such a phenotype has the possibility to cognitively ‘adapt’ to some types of events *within its own lifetime*.

4 The Nature and Mechanism of Qualia

I inferred above that the source of qualia is some sort of living substance. In this section, I consider the ontological nature of qualia and the type of event that realizes them. Whereas the arguments concerning agency were mainly analytical, with regard to qualia they are primarily scientific.

4.1 Qualia Are Energy Generated in Specialized Neurons

Qualia appear to be a form of energy. We detect them, and detection generally relies on transduction - the conversion of energy from one form to another. Qualia can carry information for both cognitive and behavioral functions, and communications theory holds that information is modulated energy. Additionally, it seems that everything recognized by modern physics is energy of some type, so feelings might be as well. This supposition is consistent with various observations. These include data from brain stimulation reward experiments, the perceived intensity of sensations as a function of neural activity, phenomenal experiences of some types, and measurements of energy consumption by the brain.

When humans undergo electrical stimulation at some brain sites, they report experiences of pleasure (Heath, 1964). So compelling is the effect on some subjects that the use of this technique raises ethical issues (Oshima and Katayama 2010). In Brain Stimulation Reward studies on rats, electrical pulses are applied to regions of a rat's brain that correspond anatomically to these human 'pleasure centers'; one that is typically targeted is the medial forebrain bundle. A rat will work for these pulses; its motivation is measured by the effort it expends to obtain the reward. AQ1

A key result from these experiments is that the strength of the reward effect depends on the total firing rate produced in the relevant neurons, not on the form of the stimulating pulse train (Gallistel et al. 1981; Shizgal, 1999; Simmons and Gallistel 1994). An explanation of this phenomenon suggests itself: the relevant neurons transform the electro-chemical energy of the neural spikes to the phenomenal energy of pleasure, the perceived intensity of the pleasure is proportional to the aggregate energy of the converted neural pulses, and a rat works harder for rewarding pleasure that is more intense.

The strength of sensory qualia also appears to depend on the energy in the associated neurons. In a study on odors, the experienced intensity of a smell correlated with the rate of neural impulses in the amygdala (Winston et al. 2005). Similarly, Mather (2006) reports that "the most successful model of loudness perception … proposes that the overall loudness of a given sound is proportional to the total neural activity evoked by it in the auditory nerve." The rate of neural firings has also been observed to have considerable influence on the visual perception of brightness (Kinoshita and Komatsu 2001) and on the tactile perception of the amplitude of a surface vibration (Bensmaia 2008).

Certain types of phenomenal experiences also support the hypothesis that qualia result from an energy transduction, rather than information processing. A strong blow to the head produces the visual sensation of 'seeing stars'. Apparently, some of the mechanical energy of the jolt is transduced to action potentials in those neurons that convert the pulses to visual qualia. Also, visible and audible white noise carry no information, so there can be no symbolic representation in the resulting neural activity to the effect that 'this is noise'. Yet, an experience of such a phenomenon provides us with knowledge of its random character and its strength; how can this be? Although noise lacks information, it does consist of energy. Evidently, the energy comprising the neural noise is converted to a phenomenal experience, one which retains the relative intensity and spectral properties of the aggregated neural pulses.

Additionally, some measurements of energy consumption by the brain support the hypothesis that a portion of that energy is converted to feelings. Using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), Raichle (2006) measured the brain's responses to controlled stimuli (in terms of changes in blood flow). Because the increase in energy consumption due to the stimuli was much less than expected, he surmised that "the brain apparently uses most of its energy for functions unaccounted for." Raichle calls this 'dark energy' and posits that it supports intrinsic neural activity for functions like the maintenance of information.

I hypothesize that at least part of this energy is spent in the production of qualia, a fundamental form of energy that was not captured in Raichle's measurements. When awake, we are continuously subjected to feelings of various kinds (both conscious and subconscious) in sensing our external and internal environments. If phenomenal experiences are a form of energy, generating those feelings would increase the baseline metabolic rate of the brain. This would explain the considerable amount of energy that was 'missing' in Raichle's studies.

In principle, the sort of experiment performed by Raichle could provide a means for falsifying the hypothesis that qualia are a form of energy. If that theory is correct, measurements of regions that are sources of feelings should show more 'missing' energy than locations that are not. Such an experiment depends on identifying the areas of the brain that produce qualia, and on a measurement technique with sufficient spatial resolution to distinguish those regions from locations that don't generate feelings.

4.2 A Phenomenal Experience Depends on a Local Event

There are two alternatives regarding the spatio-temporal nature of the event(s) that cause a phenomenal experience. One is the computational hypothesis that a quale results from a causal pattern, such that the existence and character of a quale depend on multiple events distributed over space and time. The other possibility is that a quale is caused by a singular event at a particular space-time location. For the latter alternative, a quale must depend on the type of matter and/or energy at its location; otherwise, it would be under-determined. From the concept of an action, I inferred above that a phenomenal experience has its origin in some type of substance. Here, I argue again for this claim - this time mainly from science.

The causal-pattern hypothesis faces a challenge from physics in the principle of locality, which holds that an event at a space-time location depends only on what is at that location. Einstein expressed the importance of this principle in a letter to Max Born (1971, 171): "If this axiom were to be completely abolished, the idea of the existence of quasi-enclosed systems, and thereby the postulation of laws which can be checked empirically in the accepted sense, would become impossible." Intuitively, the locality principle seems correct; how could an event at some instant be influenced by things that are not at the event's location at that instant? Locality does not preclude the existence of causal 'chains' over space and time, but it does entail that the *type* of event that occurs at a time and place depends only on what exists then and there. The motion of a billiard ball may have its *historical* cause in a complex pattern of collisions involving many other balls, but the *type* of motion a ball exhibits is due only to the way it is impacted by the last ball in the sequence. In general, the locality principle is evident in laws of physics, which do not include any time delays or spatial separations between causes and effects. A changing magnetic field produces an electric field when and where the change in the magnetic field occurs. A mass is accelerated by gravity in proportion to the strength of the gravitational field at the space-time location of the mass.

Turing's (1950) canonical characterization of a computer also conforms to the locality principle. The next state of his machine depends only on the current state and the input to that state, as reflected in the computer's 'machine table'. Accordingly, an effect that depended on a pattern of prior machine states would not be the result of a computation. Furthermore, the supervenience formulation of 'minimal physicalism' as described by Kim (1998) also reflects locality: "Mental properties supervene on physical properties, in that necessarily, for any mental property *M*, if anything has *M* at time *t*, there exists a physical base (or subvenient) property *P* such that it has *P* at *t*, and necessarily anything that has *P* at a time has *M* at that time." [my underline] This precludes a 'physicalist' view of mental states as realizations of causal patterns.

It might be contended that locality does not apply to some types of events, those that exhibit what is called 'quantum entanglement.' Einstein was sceptical of this phenomenon, which he termed 'spooky action at a distance'. But this effect has been confirmed experimentally; the spin-polarizations of electrons generated in pairs and then separated seem, when measured, to influence each other instantaneously across space. The specific basis for this dependence is debated, but the possible explanations all appear to entail a non-local influence of some kind (Yanofsky, 2013).

Nevertheless, quantum entanglement can't rescue causal-pattern theories of qualia - especially if feelings are a form of energy. Information can't be conveyed using entangled properties, and there is no evidence that neural activity in the brain depends on non-local effects. Whether a particular neuron fires is fully explained by local events at its synapse; it does not depend on the history of those events. Furthermore, if a feeling is realized by an energy transform, that event must be localized - or fundamental laws of physics would be violated. Specifically, if the emerging energy-type did not come into being at the same *time* that the prior type is extinguished, there would be a violation of the conservation of energy in the interim. Or, if the new form of energy did not arise at the same *place* as the prior type, the relativistic limit on the speed of signalling would be breached.

Ironically, the physicalist view that a quale depends on a causal pattern implies some sort of non-physical causation. Consider two computers that are in qualitatively identical physical states at some instant. The first has executed the computational algorithm that is thought to be necessary for realizing some feeling, while the second has simply been placed in the same, resulting state. If the first computer has a phenomenal experience while the second does not, that difference could only be due to some non-physical influence because, by stipulation, the two computers are physically identical. Any 'memories' of the computational sequence that exist in the first computer would also be duplicated in the second - unless those 'memories' were non-physical.

Therefore, a feeling depends only on that which exists at its space-time location, which entails that qualia are determined by a particular kind of 'stuff'. In Sect. 3.3 above, I deduced that qualia originate in living matter, but it remains to consider the sort of mechanism by which they are realized, and why they are private.

4.3 Qualia Are Subjective Because They Are Spent as They Are Produced

Subjectivity concerns the process of observation. In general, observation is a form of communication in which energy of some type carries information from the event of interest, to the observer. Objective observation requires that, in principle, any observer could have received the very same modulated energy. So, for a feeling to be objectively observable *qua feeling*, some of the phenomenal energy would have to leave its source. Evidently, this is not possible for qualia. A similar circumstance exists in the cosmological phenomenon of a black hole. Any light produced by, or within the vicinity of, the 'hole' is not observable because it can't escape the gravitational force of the collapsed star.

This suggests that a feeling is subjective because it does not escape its origin. Unlike the energetic property of thermal heat, phenomenal energy is apparently not transferrable by contact, nor is it radiated. But the energy of a phenomenal experience can't simply disappear when the experience stops; it must be transformed to energy of another kind. I posit that qualia are converted to another type of energy as, and where, they are generated. The homeostatic character of life offers a clue to the nature of that energy transduction.

4.4 Qualia Are a Defense Mechanism of Their Living Micro-source

I have argued that the source of qualia is some type of living matter. In addition to metabolizing energy and materials to keep itself going, a living cell defends itself against some dangers to its well-being. One such mechanism is its construction of heat-shock proteins when the cell is confronted with various threats - like thermal changes, oxidative stress, or some toxic substances (Richter et al. 2010). I hypothesize that, like the production of heat-shock proteins, the generation of qualia serves a defensive, homeostatic function for the living matter that produces feelings in specialized neurons. I shall use the term "q-source" for this substance. I posit that action potentials in these neurons threaten the biological integrity of the q-source, and it avoids harm from the neural spikes *by converting them to feelings*.

Why does this make feelings private? The second law of thermodynamics dictates that preventing an increase in the entropy of the q-source requires the expenditure of energy, just as a refrigerator must use energy to slow the increase in the entropy of its contents. If the act of transforming neural pulses into qualia averts a threat to the biological integrity of the source of the qualia, energy must be consumed for that benefit. That energy apparently comes from the qualia themselves; if so, they never leave their source. As they are generated, qualia are transformed immediately to another type of energy; this precludes objective observation of them. I earlier analogized the q-source to a refrigerator. If the above account of qualia's generation is correct, the q-source is a remarkable sort of refrigerator. Unlike the kind of machine we use to preserve food, which requires energy from an external source, the energy used by the q-source to 'cool' the 'hot' things inside it (i.e., the action potentials) comes from those very things!

A different perspective might clarify this postulated mechanism. Living matter contains potential energy that resides in its structure, an organization of atoms and molecules that enables the substance to perform the activities that keep it alive. I hypothesize that action potentials can damage the organizational structure of the q-source. As neural activity begins to have this effect, the q-source reacts by converting the electro-chemical pulses to feelings. Ridding itself of the neural spikes in this way enables the q-source to return to its original structure, which benefit is compensated by the immediate expenditure of the qualia. Energy is thereby conserved; the energy in the neural pulses is converted to the energy of the feeling, which is instantly exchanged for the potential energy of the q-source. Accordingly, the generation of feelings is an action of the q-source that is due to its homeostatic nature; qualia are produced by living matter of some kind.

5 Concluding Remarks

The biological theory described in this paper is more scientifically conservative than the dominant, computationalist hypothesis because, while it posits a new form of energy, it does not violate any law of physics. And it explains more.

Regarding agency, the multiple realizability of a computer entails that the form of its movement has an external source (its designer), while the intrinsic nature of living matter bestows it with self-determined behavior for a self-determined goal: itself. The computational theory does not fundamentally distinguish actions from reflexes. On the biological hypothesis, actions exhibit the ‘circular’, homeostatic movement of a self-sustaining substance, while reflexes have a ‘linear’, programmed form. Functionalist theories struggle to find a causal role for the experiential aspect of a feeling, but this is not a problem for the biological theory wherein affective qualia serve as control signals in the regulatory processes by which a living organism maintains itself.

The central assumption of the computationalist view, that a phenomenal experience is determined by a causal pattern, contradicts the locality principle of physics. It thereby entails a radical form of causation that defies space and time. The biological theory does not violate locality; it postulates that a quale is the product of a singular, localized event: an energy transduction. No scientific account of subjectivity is provided by the orthodox, functionalist theory, while subjectivity is nomologically necessitated if, at the micro-level, qualia prevent an increase in the entropy of their source – a function that accords with the homeostatic character of life. Neither theory accounts for the experiential property of a feeling, but this epistemological failing is consistent with the inference that qualia are a *fundamental* form of energy.

No part of the biological theory is ad hoc. As pictured in Fig. 1, it provides integrated, mutually supporting accounts of agency, qualia and their subjectivity - all scientifically based in the thermodynamically unique, self-maintaining nature of living matter.

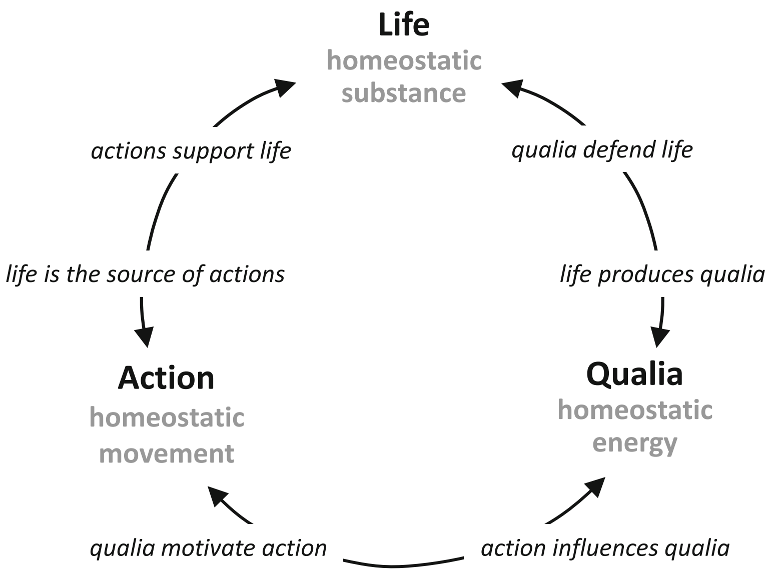

Fig. 1. Agency and qualia depend on the homeostatic nature of life.

Acknowledgments. I am grateful to reviewers Paul Schweizer and Chuanfei Chin, whose thoughtful comments and suggestions on the initial version of the paper prompted me to improve it in a number of ways. Vincent Müller also has my sincere thanks for organizing a stimulating conference, and for his help as Editor.

References

Apostle, H.G.: Aristotle's On the Soul, 416b18–19. The Peripatetic Press, Grinnell (1981). Translated by H. Apostle

Ashby, W.R.: An Introduction to Cybernetics. Martino Publishing, Mansfield Centre (1956/2015)

Barandarian, X., Di Paolo, E., Rohde, M.: Defining agency: individuality, normativity, asymmetry, and spatio-temporality in action. Adapt. Behav. **17**(5), 370 (2009)

Bensmaia, S.: Tactical intensity and population codes. Behav. Brain Rese. **190**(2), 165–173 (2008)

Block, N.: On a confusion about a function of consciousness. Behav. Brain Sci. **18**, 2 (1995)

Born, M.: The Born-Einstein letters, p. 171. Walker and Company, New York (1971)

Cannon, W.B.: The Wisdom of the Body. W. W. Norton & Company, New York (1932/1963)

Carver, C.S.: A cybernetic model of self-attention processes. J. Pers. Soc. Psychol. **50**, 1216–1221 (1979)

Carver, C.S., Scheier, M.F.: Attention and Self-regulation: A Control-Theory Approach to Human Behaviour. Springer, New York (1981)

Damasio, A.: Self Comes to Mind, p. 53. Random House, New York (2012)

Gallistel, C., Shizgal, P., Yeomans, J.: A portrait of the substrate for self- stimulation. Psychol. Rev. **88**, 228–273 (1981)

Gribbin, J.: Science: a History 1543-2001. BCA/The Penguin Press, London (2002)

Heath, R.G.: The Role of Pleasure in Behavior. Harper & Row, New York (1964)

Hume, D.A.: Treatise of Human Nature, p. 266. Oxford University Press, Oxford (1739/2000). Ed. D. F. Norton and M. J. Norton

Johnson-Laird, P.: The Computer and the Mind. Harvard University Press, Cambridge (1988)

Jonas, H.: The Phenomenon of Life, p. 79. Northwestern University Press, Evanston (2001)

Kim, J.: Mind in a Physical World. MIT Press, Cambridg (1998)

Kim, J.: Physicalism, or Something Near Enough. Princeton University Press, Princeton (2005)

Kinoshita, M., Komatsu, H.: Neural representation of the luminance and brightness of a uniform surface in the macaque primary visual cortex. J. Neurophysiol. **86**, 2559 (2001)

Machamer, P., Darden, L., Craver, C.: Thinking about mechanisms. Philos. Sci. **67**, 1–25 (2000)

MacKay, D.M.: Cerebral organization and the conscious control of action. In: Eccles, J.C. (ed.) Brain and Conscious Experience, pp. 422–445. Springer, Heidelberg (1966)

Marken, R.: Looking at behavior through control theory glasses. Rev. Gen. Psychol. **6**(3), 260–270 (2002)

Mather, G.: Foundations of Perception, pp. 116–141. Psychology Press Ltd., New York (2006)

Maturana, H., Varela, F.: Autopoiesis and Cognition, p. 79. D. Reidel Publishing Company, London (1980)

Nagasako, E., Oaklander, A., Dworkin, R.: Congenital insensitivity to pain: an update. Pain **101**, 213–219 (2003)

Oshima, H., Katayama, Y.: Neuro-ethics of deep brain stimulation for mental disorders: brain stimulation reward in humans. Neurol. Medico-Chirurgica **50**, 845–852 (2010)

Powers, W.T.: Behavior: The Control of Perception. Aldine, Chicago (1973)

Prigogine, I.: Time, structure and fluctuations. Science **201**, 777–785 (1978)

Pross, A.: The driving force for life's emergence: kinetic and thermodynamic considerations. J. Theor. Biol. **220**, 393–406 (2003)

Pross, A.: What Is Life? Oxford University Press, Oxford (2012)

Raichle, M.: The brain's dark energy. Science **314**, 1249–1250 (2006)

Richter, K., Haslbeck, M., Buchner, J.: The heat shock response: life on the verge of death. Mol. Cell **40**(2), 253–266 (2010)

Russell, B.: The Problems of Philosophy. Oxford University Press, London (1959)

Sachs, J.: Aristotle's On the Soul and On Memory and Recollection. Green Lion Press, Santa Fe (2001). Translated by J. Sachs

Schneider, E.D., Kay, J.J.: Life as a manifestation of the second law of thermodynamics. Math. Comput. Model. **19**(6-8), 36 (1994)

Schrödinger, E.: What is life? With Mind and Matter, p. 79. Cambridge University Press, Cambridge (1944/1967) AQ2

Schulze, G., Mariano, M.: Mechanisms of Motivation, vol. 15, pp. 1–5. Trafford Publishing, Victoria (2004)

Shizgal, P.: On the neural computation of utility: implications of studies from brain stimulation reward. In: Kahneman, D., Diener, E., Schwartz, N. (eds.) Well-being: the Foundation of Hedonic Psychology, pp. 500–524. Russel Sage Foundation, New York (1999)

Simmons, J.M., Gallistel, C.R.: Saturation of subjective reward magnitude as a function of current and pulse frequency. Behav. Neurosci. **108**, 151–160 (1994)

Spencer, H.: The Principles of Psychology, p. 593. Longman, Brown, Green, and Longmans, London (1855)

Turing, A.M.: Computing machinery and intelligence. Mind **59**, 433–460 (1950)

Winston, J., Gottfried, J., Kilner, J., Dolan, R.: Integrated neural representations of odor intensity and affective valence in human amygdala. J. Neurosci. **25**(39), 8903–8907 (2005)

Yanofsky, N.: The Outer Limits of Reason. MIT Press, Cambridge (2013)

Author Query Form

Book ID :	**469986_1_En**
Chapter No.:	**6**

Please ensure you fill out your response to the queries raised below and return this form along with your corrections

Dear Author
During the process of typesetting your chapter, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the 'Author's response' area provided below

Query Refs.	Details Required	Author's Response
AQ1	The citation 'Oshima et al. (2010)' has been changed to 'Oshima and Katayama (2010)' to match the author name in the reference list. Please check here and in subsequent occurrences, and correct if necessary.	
AQ2	Please check and confirm the edit made in Ref. 'Schrödinger (1944/1967)'.	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	• • • under matter to remain	
Insert in text the matter indicated in the margin		New matter followed by or
Delete	/ through single character, rule or underline or through all characters to be deleted	or
Substitute character or substitute part of one or more word(s)	/ through letter or through characters	new character / or new characters /
Change to italics	under matter to be changed	
Change to capitals	under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Change bold to non-bold type	(As above)	
Insert 'superior' character	/ through character or where required	or under character e.g. or
Insert 'inferior' character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	
Insert single quotation marks	(As above)	or and/or or
Insert double quotation marks	(As above)	or and/or or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	/ through character or where required	
Reduce space between characters or words	between characters or words affected	