2020

Consciousness as a factor in evolution

Kenneth A. Augustyn
Michigan Technological University, kaaugust@mtu.edu

Consciousness as a factor in evolution

Kenneth A. Augustyn

Department of Physics
Michigan Technological University, Houghton Michigan, 49931 USA
kaaugust@mtu.edu

Abstract

What I call *the mind* began as a non-conscious robotic biochemical process control system in the very earliest forms of life. As life evolved, problems in control became more difficult and exceeded the computational capabilities of the organisms. Nature discovered a means of transcending computable physical processes resulting in non-computational subjective mental capabilities that, while still not conscious, had a degree of genuine autonomy from the physical world. These autonomous subjective wants and goals now affected the course of (but not the mechanism of) evolution. The integrated amalgam of robotic and transrobotic unconscious capabilities eventually gave rise to consciousness, which became an even more important factor in the course of evolution.

The processes responsible for transrobotic mentality are conjectured to leave evidence in the physical world in the form of violations of conservation laws, evidence that future experiments may be able to detect.

Keywords: *Biological mentality, Mind-matter, Consciousness, Robotic mentality, Transrobotic mentality, Evolution*

Manuscript received 20 February, 2020; Revised version received 27 March 2020; Accepted 5 May 2020
Journal of Cognitive Science 21-3:489-505, 2020

1. Introduction

This paper informally reviews and expands upon a theory of biological mentality first presented in (Augustyn, 2019), a theory that has been reinforced by many discussions during the Third Workshop on Biological Mentality (Augustyn, 2020) and elsewhere. The theory includes conjectured biological mentality processes analogous to parametric down conversion, predicts local violations in conservation laws, and is thereby experimentally falsifiable.

Mind-body theories have been discussed for centuries and there are whole classes of such theories that I reject as inadequate. It would be a divergence from the purpose of this paper to delve into these theories. I only mention them at all so that the reader can see in advance where I will not be going.

I reject theories that assert that consciousness has no genuine independence from the physical world, that it is some kind of illusion, that it is the byproduct of computing, that it is entirely robotic.

I reject theories that claim the physical world is an illusion because it is only consciousness that exists.

I reject theories that claim that *everything* in the universe is conscious, that consciousness is present in every piece of the physical world, that every rock and every atom has its own kind of consciousness.

I reject theories that connect consciousness with the physical world *directly*, without considering the unconscious.

And I reject all theories with no possibility of experimental confirmation or refutation.

2. Conscious and Unconscious

I begin with a notional overview of how I use terms such as consciousness and *unconscious* without giving any formal definitions. I use the term

biological mentality to encompass both conscious and unconscious capabilities (or just the unconscious capabilities of organisms lacking consciousness). Consciousness has an incoming side, *awareness*, and an outgoing side, *free will.* Both will now be discussed.

2.1 Conscious Awareness

I am aware of the physical world around me, how things in this world look, taste, feel, smell, and also of internal sensations such as hunger and thirst. This awareness is obviously limited. I cannot without instrumentation sense the presence of microorganisms or the signals broadcast from radio stations. I have no direct sensation what my liver or pancreas is doing. And each physical object that I am able to sense has much deeper levels of reality that I cannot sense directly, e.g. the atoms making up my desk or the fusion reactions taking place in the sun.

I am also aware of thoughts, inner speech, emotions, memories, and sense of myself as an entity.

Awareness comes to me passively. I don't have to make an effort to be aware, e.g. I don't have to *try* to taste the apple I am biting into. (I may have to try hard to focus my attention on something that I *want* to be aware of, such as searching for my misplaced glasses.)

The content of awareness is personal and idiosyncratic. I see red and green as very different colors, and have no way of imagining how red and green look to someone who cannot distinguish between them.

2.2 Conscious Free Will

Consciousness also has an outgoing side that I call *free will.* Free will refers to those thoughts and actions not predetermined by physical law, not arbitrary or random, not the result of any kind of compulsion. Free will has two components: *free choice* is the power of choosing without the constraint of necessity or fate, and *volition* is the power to carry out a free

choice.

I can freely choose what to do next with my body and I can usually execute these choices (within limits of course). I can also choose what to think about, what to silently talk to myself about, what songs to "play in my head" - again within limitations. I may freely choose to do something only to find that I cannot do it. For example, I may encounter a person that I know very well but find myself unable to come up with his name even though I want to greet him. Hence I distinguish between the free choice of an action and the actual execution of that choice (volition). An action that I did not choose, such as a muscle spasm, is not an act of volition.

Unlike awareness, which seems to come to me without effort, it takes obvious effort for me to make a choice or to take action.

2.3 Unconscious

Consciousness, both on the incoming side (awareness) and the outgoing side (free choice and volition) interfaces with the physical world exclusively by means of the unconscious which is much more than a biological computer.

On the incoming side of consciousness (awareness), we have the feeling that we sense the outside world directly but we do not. The content of our awareness is produced and delivered to consciousness by unconscious processes. For example, two areas of identical objective color (i.e., identical pixel values) will be consciously perceived as two different colors because of context as illustrated in a demonstration by Donald Hoffman on YouTube[1]. Technically this is an illusion, but we have to be very careful here! *Illusion* is a loaded word. Look at any dictionary definition of illusion and you will find words like "false", "wrongly perceived", "deceptive". The subjectively perceived colors constructed by unconscious processes that take context into account are not deceptive. The unconscious had

1 https://www.youtube.com/watch?v=oadgHhdgRkI Start at 5:00

"good reasons" for presenting the colors as different because of the partial shadowing.

On the outgoing side of consciousness, free choice and volition also depend on unconscious processes. As an analogy, consider driving an automobile. You, the driver, are in control. But do you decide which cylinder will fire next? How much the spark advance on that cylinder should be? Or any of thousands of variables involved in the car responding to your high-level control? Of course not. These are all handled by physical control systems engineered and installed by the manufacturer. You, the driver, don't even have to know that there is such a thing as spark advance that must be controlled. Your ultimate control of the car is achieved via the high-level variables that you do control, such as the position of gas pedal.

The control systems inside an automobile are *robotic*. They can be fully described by computable functions. Even if a digital computer is not used for implementation (as was the case for automobiles years ago that used spinning weights and springs to control spark advance) the control function is still computable.

What is different between driving a car and "driving yourself" is that all of the car's internal control systems are entirely robotic whereas your unconscious has both robotic and *transrobotic* capabilities. The idea of transrobotic capabilities will be discussed. For now, the important point is that the unconscious is not merely robotic. It is much more than computation.

3. The theory of transrobotic mentality

3.1 Evolutionary state changes

It is often assumed, explicitly or implicitly, that the unconscious functions mechanically, as does a computer. I believe this is only partially correct. There is a purely physical part that I call the *robotic* unconscious. In

addition, the unconscious also has what I will call *transrobotic* capabilities which evolved long before consciousness itself evolved.

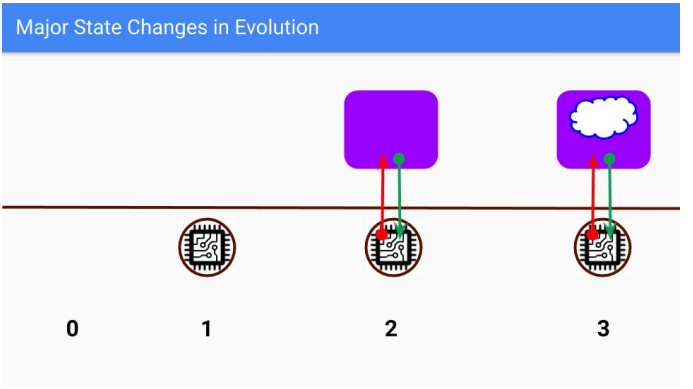

Figure 1. This figure depicts the timeline of major state changes in the evolution of life, with time proceeding from left to right starting with the big bang.

- At time 0, the physical world existed (area below the line) but life had not yet emerged.
- At time 1, life (represented by the circle above 1) emerged and was entirely *robotic*. That is, all of the control processes of all life forms existing at this time were handled by self-organizing biochemical reactions and other physical processes that are entirely computable.
- At time 2, a new evolutionary development enabled some organisms to perform *transrobotic* processes depicted by the red and green arrows connecting the body of the organism to the organism's unconscious (depicted as a purple area above the line separating the physical world from the newly-emerged nonphysical world of biological mentality), something robotic organisms do not have. Organisms at this stage are not conscious but they have capabilities that transcend computing including the capability to have unconscious phenomenal experiences

and to exercise free choice in willful actions.

- At time 3, consciousness emerges within the transrobotic unconscious foundation. All conscious awareness and conscious free choice interact with the body not directly but via the transrobotic unconscious.
- Times 0,1,2,3 are separated by billions of years and so unconscious biological mentality had billions of years to develop before consciousness emerged.

In today's world, we have living organisms representative of each stage of evolution. Perhaps bacteria are still stuck in the robotic level of development, single-cell eukaryotes have more than robotic capabilities but no consciousness, while birds, mammals, and cephalopods have various degrees of consciousness.

3.2 Imagining organisms without consciousness but still more than robotic

It is difficult to imagine what it might be like to have a transrobotic unconscious but not consciousness itself. We can imagine our own unconscious still functioning at times we are not conscious. Sleepwalking illustrates this.

> *Sleepwalking is a behavior disorder that originates during deep sleep and results in walking or performing other complex behaviors while asleep.*
>
> *Sleepwalking usually involves more than just walking during sleep. Symptoms of sleepwalking disorder range from simply sitting up in bed and looking around, to walking around the room or house, to leaving the house and even*

driving long distances[2].

3.3 Two kinds of transrobotic processes

Figure 1 depicts two kinds of transrobotic processes: a red arrow depicting energy withdrawn from the body (and hence from the physical world) and converted into units of transrobotic unconscious mentality, and a green arrow for transrobotic unconscious mentality converted back to energy injected back into the body in order to execute free choices such as willed movements and willed focus of attention. Both red and green processes violate conservation laws. Hence, unconscious mentality leaves traces in the physical world, traces that may be detected experimentally as violations of mass-energy and momentum conservation laws.

In the laboratory, a pair of entangled photons can be created from a single photon passing into a nonlinear crystal in a process known as spontaneous parametric down conversion[3]. In the emitted pair of photons so created, measurable properties such as energies, directions, and polarizations of the two photons when measured will be correlated no matter how far apart the two photons have traveled away from each other. Because of these inherent correlations, if one photon is measured to produce a definite property, that property of the other correlated photon is also determined. I conjecture that incoming transrobotic processes employ a process to down conversion for splitting quanta of energy (while outgoing transrobotic processes use a process similar to up conversion).

2 https://www.sleepfoundation.org/articles/sleepwalking

3 See any quantum optics textbook, or H. Fei et al, Phys. Rev. Letts. 78, 1679 (1997).

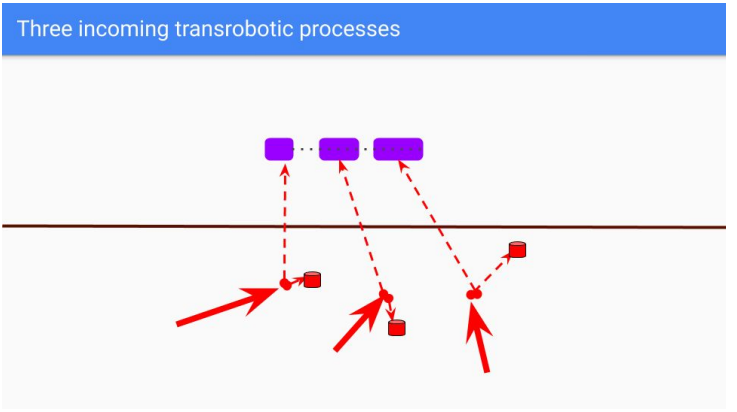

Figure 2. Incoming transrobotic processes. The three solid red arrows depict three quanta of energy in a living organism that each split into two dashed arrows, one of which stays in the organism (below the line) while the other exits the physical world and converts to a unit of unconscious transrobotic mentality (above the line). These mentality units grow and are entangled together as represented by the dotted line connecting them.

The energy of each dashed arrow remaining in the organism is less than the energy of the solid arrow that produced it because some energy was withdrawn from the physical world to power transrobotic mentality, which at this point is limited to awareness functions. This incoming transrobotic process violates mass-energy conservation since it literally withdraws energy from the physical world.

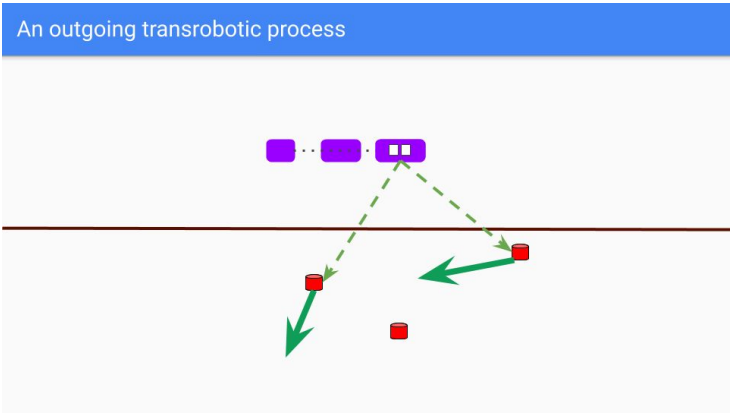

Figure 3. Outgoing transrobotic processes. Here, transrobotic unconscious mentality executes a free choice by injecting two partial units of energy (dashed green arrows) into the physical world at a time-locations coincident with the two partial units of energy (two of the three dashed red arrows of Figure 2) with which it has remained entangled. Each of the two entangled partial units of energy, one from outside and one from inside the physical word, combine to form a full unit of energy (solid green arrow) applied to change the direction of motion of physical biomolecules involved in some otherwise-robotic process.

These transrobotic processes enable a nonphysical subjective mentality that can actively interfere with at least some of the organism's otherwise-deterministic robotic processes. Events in the physical world are no longer limited to the inevitable consequence of prior events (determinism) or the purely random (as in quantum physics). Now there are subjectively willful events. Organisms at some level of evolution behave in ways that are not only uncomputable but are not even deterministic.

4. The evolution of biological mentality

In (Augustyn, 2019) I conjectured in more detail how biological mentality could have evolved in a sequence of state changes: robotic unconsciousness, then an amalgam of both transrobotic and robotic unconsciousness, then eventually consciousness. The earliest and most primitive organisms were entirely robotic. With the advent of single-cell eukaryote organisms containing internally reproducible mitochondria, an energy-availability limitation was overcome, enabling the discovery by Nature of the transrobotic processes that are part-physical and part-nonphysical. These processes exchange energy between the physical organism and its non-physical unconscious mentality. Elements of the organism's mentality space remain entangled with the organism's physical body, and via this entanglement mentality can inject energy back into the organism resulting in non-deterministic actions as the affected biomolecules deflect from their otherwise-deterministic trajectories because of the injected energy. Transrobotic processes resulted in the first instance of transrobotic mentality billions of years before anything remotely like consciousness emerged.

4.1 The scope of robotic mentality

Robotic mentality is the foundation of all biological mentality. It purely mechanical, purely deterministic (or deterministic plus random[4]). Biochemical reactions that control other biochemical reactions are examples of robotic mentality. Robotic mentality in organisms could, in principle, be precisely simulated by computer programs.

Self-organization involves the generation and maintenance of order in dissipative systems requiring a constant input of energy. Self-organization

4 By *deterministic* I mean either fully deterministic or deterministic with randomness as provided by quantum physics, such randomness providing no real escape from the fatalistic nature of determinism.

is deterministic and fully consistent with conservation laws. Vast numbers of biochemical pathways jointly form complex dynamic networks. The pathways we know about have been plotted manually[5] and more recently by automated techniques.[6] All are involved in control as well as production. That is, biochemical processes control other such processes by providing inputs, catalysts, or otherwise.

> *Cells, with their constant energy consumption and myriads of local interactions between distinct proteins, lipids, carbohydrates and nucleic acids, represent the perfect playground for self-organization. It therefore comes as no surprise that many properties and features of self-organized systems, such as spontaneous formation of patterns, nonlinear coupling of reactions, bi-stable switches, waves and oscillations, are found in all aspects of modern cell biology. Ultimately, self-organization lies at the heart of the robustness and adaptability found in cellular and organismal organization, and hence constitutes a fundamental basis for natural selection and evolution.* (Wedlich-Söldner, 2018)

Robotic mentality evolved in complexity and capability as life evolved. While it is not all there is to a more advanced biological mentality, it is the physical infrastructure upon which higher levels of transrobotic mentality operate.

4.2 The emergence of transrobotic mentality

Robotic mentality requires onboard computational resources. Just as

5 https://www.roche.com/sustainability/philanthropy/science_education/pathways.htm

6 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6466808/

the automobiles of years past used analog devices for internal control functions (such as spinning springs and weights for spark advance control), early organisms used internal analog resources such as self-organizing biochemical reactions as "devices" for control purposes. We can quantify the computational capability of an analog device as-if it were implemented by a digital computer. Hence we can treat the robotic mentality of an organism as a hierarchy of digital control systems, each system executing on a digital computer of limited capability.

The underlying premise of transrobotic mentality is that fully-robotic organisms of the distant past encountered problems of control that were beyond their computational capabilities. Some such problems may have been solved by the evolutionary emergence of additional computational capabilities, giving the more-computationally-powerful organisms a selective advantage[7]. However, in other cases of control failure, evolution discovered processes in Nature that broke through the physical world and enabled the organism to disable, reset, or otherwise disrupt the deterministic flow of events in that organism's failing situation. Once discovered, Nature found other ways of exploiting these processes to override robotic control, giving such organisms additional selective advantage. Unconscious biological mentality evolved into an amalgam of interlinked robotic and transrobotic processes, endowing the more advanced organisms with subjective "mental powers", powers with genuine independence from the physical world. Consciousness emerged within this unconscious amalgam and, as mentioned before, is dependent on the unconscious to provide an adequate liaison to the physical world as well as to provide deep mental powers that consciousness often takes credit for (as when a consciously-worked problem is put aside, only to have the solution delivered to

7 As an analogy, the introduction of electronic ignition modules enabled the auto industry to solve control problems relevant to fuel economy and exhaust emissions, problems that could not be solved with spinning weights and springs.

consciousness while consciousness was focused elsewhere).

4.3 Consciousness as a factor in evolution

Physical objects, including primitive living organisms having only robotic capabilities, do not literally "want" anything no more than a ball rolling down a hill "wants" to be doing so. The term *want*, when applied to a static object or robotic process such as a computer program, is a only metaphor. Evolution by natural selection is correctly called *blind* because it too is a robotic process that does not "want" to accomplish anything.

Only with the emergence of a transrobotic mentality does the term *want* have literal meaning and this literal meaning resides only within the organism's private nonphysical subjective world. Once organisms had a subjective mentality, a new dimension in the evolutionary equation came into being. Living organisms could now (literally) *want.* Their individual subjective *wants*, as well as other subjective emotions, influenced their objective physical behaviour, and thus affected all of evolutionary history from that point on. Blind evolution now had a new, nonphysical source of variation.

The emergence of consciousness and especially the conscious capabilities of humans often came at a high price. Conscious autonomy enables people to knowingly and willfully contradict the biological imperative to survive, e.g., many people knowingly chose probable suicide in exchange for temporary recreational highs while many others have chosen certain suicide in the service a supposed higher cause (e.g., signing up to be a kamikaze pilot or suicide bomber). It also enables people to knowingly and willfully circumvent biological imperatives to reproduce.

There is yet another aspect of human consciousness that, rather than contradicting the biological imperative to survive, goes far beyond this imperative. That is the strong human desire to survive after death. Whether or not one believes in life after death, the hope for it is (usually) there.

Only humans, I presume, have the mental power to comprehend the inevitability of their own personal death and to conceive of a life after death. The subjective human concept of life after death may be yet another nonphysical factor influencing the course of evolution.

For example, Charles M. Fair (Fair, 1969) posited that conscious realization of the inevitability of personal death (a realization that perhaps only humans are capable of having) triggers unconscious flight-or-fight emergency systems to go into a high-alert state, interfering with higher level mental capabilities. Whereas a sudden shock may produce a temporary "brain freeze" (e.g., a person may see a shooting, punch in 911, forget to press send, and wind up holding a useless cell phone until they calm down), a continuous state existential fear not only continuously interferes with our higher mental capabilities but also damages the physical body. Accepting the mere possibility of the very abstract concept of life after death negates the logical inevitability of death-as-final and allows the unconscious emergency systems to subside, presumably improving both survival value and mental powers. I cite this as an example of how a consciously constructed abstract concept might "act back" on the mind that embraces it and thereby result in yet another variation for natural selection.

5. Experiments

Transrobotic mentality is conjectured to leave evidence in the physical world as violations of conservation laws, evidence that is experimentally detectable in principle. As (Cucu and Pitts, 2019; Pitts 2019) have shown, energy conservation holds if there is symmetry over time. Momentum conservation holds if there is symmetry over space. If there are time-places where such symmetries fail due to extra-physical influence, conservation laws fail there and then, while holding elsewhere in the physical universe. Transrobotic processes leave a trail of evidence in the physical world as

local intermittent violations of conservation laws. Hence transrobotic mentality is an experimentally falsifiable conjecture.

That said, as of now there is no empirical evidence whatsoever that living organisms violate conservation laws. My hope is that experimentalists will study Pitts and Cucu and that in doing so they will find motivation to design and conduct tests for conservation violations in closed systems containing living organisms.

References

Augustyn, K. 2019. Physical Foundations of Biological Mentality. *Journal of Cognitive Science* 20-2:195-214.

Augustyn, K. 2020. Special Issue on the Third Workshop on Biological Mentality. *Journal of Cognitive Science* (forthcoming)

Cucu, A.C. and Pitts, J.B. 2019. How Dualists Should (Not) Respond to the Objection from Energy Conservation. *Mind and Matter* 17:1 95-121.

Fair, Charles M. 1969. *The Dying Self,* Wesleyan University Press, Middletown CT.

Pitts, J.B. 2019. Conservation Laws and the Philosophy of Mind: Opening the Black Box, Finding a Mirror. *Philosophia* (2019).

Wedlich-Söldner, R., Betz T. 2018. Self-organization: the fundament of cell biology. *Philos Trans R Soc Lond B Biol Sci.* 373(1747).