Philosophy of Science (2022), **89**, 70–88
doi:10.1017/psa.2021.15

ARTICLE

Energy Requirements Undermine Substrate Independence and Mind-Body Functionalism

Paul Thagard[†]

Philosophy Department, University of Waterloo, Waterloo, ON, Canada
Email: pthagard@uwaterloo.ca

(Received 09 April 2020; revised 05 August 2020; accepted 12 November 2020)

Abstract

Substrate independence and mind-body functionalism claim that thinking does not depend on any particular kind of physical implementation. But real-world information processing depends on energy, and energy depends on material substrates. Biological evidence for these claims comes from ecology and neuroscience, while computational evidence comes from neuromorphic computing and deep learning. Attention to energy requirements undermines the use of substrate independence to support claims about the feasibility of artificial intelligence, the moral standing of robots, the possibility that we may be living in a computer simulation, the plausibility of transferring minds into computers, and the autonomy of psychology from neuroscience.

Keywords: Energy; mind-body functionalism; substrate independence; ecology; neuroscience; neuromorphic computing; deep learning; information

1. Introduction

In the *Matrix* movies, people are kept in a computer-generated dream world because intelligent machines are using human bodies as a power source. This story is ridiculous because bodies are inefficient batteries, but it does suggest the importance of energy to minds and computers. Without energy, there is no thinking and no computation. Nevertheless, most reflections on mind and computation ignore the impact of energy on processing information that supports intelligence.

Unavoidable tradeoffs between energy and information undermine claims that have been made about the relations between minds, bodies, and computers, including substrate independence, mind-body functionalism, and multiple realization. Facts about energy requirements reveal flaws in the arguments that have been made for these claims. Recent advances in biology concerning the relation between information and metabolism converge with new ideas in computer science about tracking

---

[†] The inspiration of this paper was a convergence of talks by Chris Eliasmith and by Mary O'Connor at the University of Waterloo in November 2019. I am grateful to both for comments on an early draft, and to anonymous reviewers for productive suggestions.

the inference to energy ratios of computing platforms. These developments substantially reduce the plausibility of substrate independence.

1.1 Substrate independence, functionalism, and multiple realizability

Substrate independence is the hypothesis that mind and computation do not depend on any particular kind of physical implementation (Bostrom 2003; Tegmark 2017). Thinking and computing can potentially operate in many different substances including brains, digital computers, analog computers, and force fields in space aliens.

The term "substrate independence" is a relatively new (post-2000) characterization of a philosophical position called "functionalism" that Levin (2018) characterizes as follows:

> Functionalism is the doctrine that what makes something a thought, desire, pain (or any other type of mental state) depends not on its internal constitution, but solely on its function, or the role it plays, in the cognitive system of which it is a part. More precisely, functionalist theories take the identity of a mental state to be determined by its causal relations to sensory stimulations, other mental states, and behavior.

Functionalism originated in the 1960s in arguments of Hilary Putnam (1975) that minds can abstractly be considered as Turing machines with operations independent of whether they are performed by brains or particular kinds of computers. I prefer the term "substrate independence" to "functionalism", which has very different meanings in other fields such as sociology, psychology, education, and engineering.

Arguments for substrate independence and functionalism are often based on examples of multiple realization where similar kinds of thinking or computation occur in different physical systems (Bickle 2020; Polger and Shapiro 2016). For example, perception and inference operate in octopus brains, which are organized very differently from human brains, and Microsoft Word and other software programs run equally well in different brands of computers. Energy considerations show that multiple realizability is much more limited than suggested by thought experiments.

1.2 Why substrate independence matters

The doctrine of substrate independence has implications for many important questions in science, engineering, and philosophy. It supports the possibility of computer systems with human-level intelligence understood as abstract information processors rather than as dependent on details about brains and bodily inputs (Tegmark 2017). Full artificial intelligence appears more feasible if substrate independence is true, although it might be achievable by engineering feats that ignore how human minds work.

Substrate independence would also make it more plausible that computers could eventually have moral status that justifies attributing praise and blame to them on ethical grounds (Bostrom and Yudkowsky 2014). Computers are not yet known to be conscious, but substrate independence implies that consciousness does not depend on any special causal powers of human brains, so that AI could eventually be conscious and capable of the kinds of feelings that make people responsible for their actions.

Substrate independence is also an undefended premise in arguments that we may now be living in a computer simulation (Bostrom 2003; Greene 2018). If powerful computers could fully simulate these human minds in ways not determined by how brains work, it is conceivable that what we take to be our own daily experiences are just simulations running on some big computers.

Substrate independence would also increase the viability of allowing people to survive death by uploading their minds into digital computers (Blackford and Broderick 2014; Kurzweil 2005, 2012; Schneider 2019). If your thoughts are just abstract computations not dependent on the details of how your brain works, then immortality beckons.

Substrate independence also offers a general strategy for cognitive science, the interdisciplinary investigation of mind and intelligence. In contrast to the increasing influence of cognitive neuroscience in many areas of psychology and philosophy, cognitive science could operate with more abstract computational considerations independent of the details of brains that use neurons, neurotransmitters, hormones, and other biological mechanisms. The importance of these questions about brains, artificial intelligence, computer simulation, and the relation between psychology and neuroscience demand assessment of the plausibility of substrate independence.

1.3 Energy

In ordinary language, energy is vaguely the capacity for vigorous activity, but I will use the technical definition from physics. Energy is the capacity for doing work, which is defined as force times distance moved. Energy comes in different forms including kinetic (motion), potential (stored), and thermal (molecular motion), all of which are capacities for applying forces over distances. Power is the energy or work per time interval, typically measured in watts, which are joules per second, where a joule is a measure of force through a distance.

The operations of brains and computers depend on mechanisms, which are combinations of connected parts (entities) whose interactions (activities) produce regular changes that accomplish results (Bechtel 2008; Craver and Darden 2013; Glennan 2017; Thagard 2019b). All mechanisms require energy to carry out the required interactions. In brains, the main parts are neurons that use glucose as their primary energy source to carry out changes such as the firing of one neuron exciting or inhibiting the firing of other neurons. In contrast, the main parts of computers are silicon chips connected by wires using electrons to process information within and between chips. Substrate independence assumes that there is no important difference between chemical energy in brains and electrical energy in computers.

2. Why energy matters

The Turing machine is not a mechanism, but rather a mathematical model in the form of an abstract device that manipulates symbols on a strip of tape according to rules. The unreality of Turing machines is evident from the infinite length of the tape and the absence of a power source for the manipulation of symbols. Turing machines operate independently of the bounds of time, space, and energy.

In contrast, physical computers and brains are bounded in time and space and are subject to energy limitations. Real systems take inputs from their environments, process information about those inputs, and act on the environments by outputs such as communication and physical actions. Energy usage in cognitive systems has five important characteristics concerning limits, costs, consumption, efficiency, and tipping points.

1. *Energy is limited.* For all systems, whether electrical or biochemical, environments provide only a finite amount of energy to fuel their operations. Biological systems are limited by getting enough food to provide the glucose that makes neurons function. Electrical systems are limited by the availability and cost of electricity.
2. *Acquiring energy has costs.* Even when energy is available in the environment, there are costs associated with acquiring it. Biological systems spend time and energy in gathering and storing food. Owners of electrical systems must pay for electricity and the infrastructure that delivers it such as generating stations and transmission lines.
3. *Input, output, and processing of information expend energy.* Mechanisms consume energy to sense environments by supporting biological organs such as eyes and ears or by supporting electrical sensors such as cameras and microphones. Processing information acquired by the senses also requires energy in forms such as the glucose used by neurons or the electricity used by digital computers. Outputs in the form of actions on the environment can be particularly expensive because of the glucose required to move muscles and the electricity required to control motors in robots.
4. *Energy must be used efficiently.* Because gaining energy is costly, cognitive systems must ensure that they get a good ratio of production to energy. Biological systems need to balance their acquired food against amounts of perception, cognition, and action, and electrical systems also need to balance electricity usage against amounts of sensing, processing, and moving. Relevant measures of energy efficiency include perceptions/calories in biological systems and inferences/joules in electrical systems, where one calorie equals approximately 4.18 joules. According to Lane and Martin (2010) and Szathmáry (2015), a major evolutionary transition occurred when the acquisition of mitochondria allowed cells to have more energy per gene.
5. *Qualitative tipping points are energy-sensitive.* Quantitative changes in energy use can tip into qualitative changes when critical thresholds are passed. For example, if a biological organism drops below a quantity of energy acquisition needed to sustain its body, it will undergo the qualitative change of death by starvation. If an electrical system uses unsupportable quantities of electricity, it will undergo the qualitative change of being discarded by its users.

Let us now examine in more detail how these five characteristics operate in biological and computational systems. Here is a sketch of my general argument:

1. Real-world information processing depends on energy.
2. Energy depends on material substrates.
3. Therefore, information processing depends on material substrates.
4. Therefore, substrate independence is false.

This argument is not strictly deductive as “depends on” may not be transitive. The sketch needs to be fleshed out by considering evidence for premises 1 and 2 in biological and computational systems. My conclusion is not that substrate independence is always false, just that there are important real-world cases where it fails.

3. Energy and information in biology

Different forms of energy are convertible into each other, for example, when the chemical energy in gasoline is converted into mechanical energy to move a car. But conversion is subject to the five characteristics of limits, costs, efficiency, consumption, and tipping points. Using gasoline to power cars will not work well if there is little gasoline extractible from the environment, the cost is too high, consumption is excessive or inefficient, or there are tipping points leading to qualitative failures. These problems are evident, most generally, in the field of ecology and, more specifically, in neuroscience.

3.1 Ecology

Whereas the cognitive sciences have long attended to information processing while ignoring energy, the field of ecology has long attended to metabolism while neglecting information. O'Connor et al. (2019, 3) reorient ecology by proposing five principles that integrate information into ecological understanding.

Principle 1: Information is a fundamental feature of living systems, and therefore, also of all ecological systems.

Principle 2: Syntactic and semiotic information interact in feedbacks, with energetic processes and material cycles, to influence structure, function and organization in ecological systems.

Principle 3: Information processing requires energy and materials; therefore, supply of energy and materials and thermodynamic constraints can limit information processing

Principle 4: Information processing allows components of living systems to measure the environment and their own state and to measure the relationship between their state and past and expected environments.

Principle 5: Information processing systems are linked within and across scales of biological organization.

Principles 2 and 3 are the most relevant to energy and substrate independence. O'Connor et al. use the terms "syntactic" and "semiotic" as follows:

> Syntactic information exists in any spatial or temporal arrangement of events or objects, including the species or functional diversity of a set of interacting species … the notes and rhythms in a bird's song, or temporal pattern of sunrise and sunset. (2019, 3)

> Information contained in structure, reflecting the structure's history, can (but does not need to) represent signs or symbols that convey meaning as interpreted by an observer (semantic information …). Semiotic information is the content and the quality of semantic information as it is carried by signs (2019, 3–4).

Both kinds of information relate to energy. Syntactic information in the form of nonrandom organization is made possible by work done with energy in the environment.

Spatial and temporal arrangements such as bird songs require energy expenditures to produce them. For semantic information, the acquisition of meaning through perceptual-motor interactions with the world requires energy in the form of light, sound, or muscular motion.

O'Connor et al. (2019) stress the interdependence of information with energy and material systems. Energy is required to create, maintain, and process information. Thus, the new emphasis on information in ecological theory ties it closely with considerations about energy, providing support for my premise that information processing is dependent on energy. The kinds of energy used in biological systems such as sense organs and muscles is dependent on their material constituents. For example, hearing depends on material structures of the outer, middle, and inner ear that gain their energy from glucose fueling for cells. Movement depends on similar fueling for muscle cells.

Metabolism is the sum of chemical reactions that sustain life, including conversion of food to energy, the conversion of food and fuel to structures such as proteins, and the elimination of wastes. The principles of O'Connor et al. (2019) indicate the relevance of information to metabolism because organisms use representations with syntactic and semantic properties to help them in finding food that fuels metabolism. Energy and information are interdependent in that an organism needs energy to acquire and use information, but it also uses information to acquire energy by gathering food. All organisms and hence the ecologies in which they operate are therefore subject to balances between energy and information. Information is not just structured data, code, or text, but must also be stored, manipulated, sent, and received in processes that require energy.

3.2 Brains

Brains are often considered to be information processors, but they are also notable as energy users. A 3-pound brain in a 150-pound body has only 2% of its weight but uses around 20% of its energy (Raichle and Gusnard 2002). The brain has around 86 billion neurons whose operation requires continuous supplies of glucose. Billions of neurons build up electrical charges and send electrochemical signals to approximately 10,000 neurons often more than 100 times per second. The expended energy in each neuron needs to be replaced by blood transport to neurons and glia cells that support them. When blood supplies are blocked by mechanism failures such as strokes, heart attacks, and asphyxiation, neurons die.

The correlation between brain size and intelligence is imperfect, as elephants and some species of whales have brains with more volume and neurons than humans. But more neurons usually means more intelligence, especially when the neurons operate in brain areas associated with complex problem solving such as the prefrontal cortex. Adding neurons, however, poses an evolutionary problem because of their intense energy needs. For brains to get bigger, food acquisition needs to increase despite the limitations of environmental ability, number of hours available for foraging, and energy cost of acquiring food.

Susanna Herculano-Houzel (2016) describes the steady increase in sizes of primate brains from lemurs to gorillas but notes the dramatic increase in human brain size that occurred in Home erectus around 1.5 million years ago. Following Wrangham

(2009), she attributes increased brain size to the invention of cooking, which makes meat and plants much easier to digest and energy intensive. Because cooking dramatically increased energy efficiency, humans could evolve larger brains without having to spend more time and energy searching for additional food supplies.

Quantitative increases in human brain size led to qualitative tipping points in human intelligence with emergent capabilities not found in other animals. Around 50–100,000 years ago, large human brains became capable of recursive thought that provided capabilities for language, art, and tools that make tools (Corballis 2011; Coolidge and Wynn 2018). Recursion requires the interaction of large numbers of neurons capable of representations of representations (Thagard 2021). Without the efficiency of cooking and the development of large brains, humans would not be capable of the energy/information balance that eventually made human civilization possible. Less than 15,000 years ago, the invention of agriculture provided sufficient food sources that led to dramatic increases in the number of humans as well as the development of productive practices such as writing, counting, and using complex machines.

Thus, the evolution of human intelligence has been energy-dependent in several ways. Acquisition of sufficient food was essential to evolution of larger primate brains and got a large boost from the invention of cooking. Changing energy availability through cultivating crops and domesticating animals enabled the formation of larger human groups such as cities and divisions of labor necessary for many cultural advances. Moreover, the ways in which humans use energy in information processing is heavily dependent on the particular kinds of matter that operate in human brains – the intense but efficient operation of billions of glucose-supplied neurons.

For organisms in general and for human brains in particular, these considerations show that information processing depends on energy, which depends on the material substrate. Here "depends" means that information acquisition, processing, and application are causally affected by mechanisms with particular kinds of parts, interactions, and energy sources. Without such mechanisms, information processing would not be effective and efficient.

4. Energy in computing

Because digital computers run on electricity, which is currently abundant, it might seem that energy is not an issue for computing. Energy is largely ignored in the theory of computing such as Turing machines, but its practical relevance is evident in the development of neuromorphic computers and machine learning.

4.1 Neuromorphic computing

Carver Mead (1990) proposed the development of neuromorphic (brain-like) computers to deal with two potential limitations in computing technology: speed and power consumption. The speed of computers has increased exponentially since the 1960s in accord with Moore's law that the number of transistors on a chip doubles every two years. But physical limits such as problems of heat dispersion put limits on how much smaller transistors can be built, so computer engineers must look to parallelism as an alternative way of increasing computers' speed.

Most computers today are built with multiple "cores," which allow parallel processing, but the number of cores is usually small, e.g., 4 or 28. In contrast, the brain has 86 billion neurons operating in parallel, which allows it to outperform current computers whose chips do trillions of operations per second in contrast to the approximately 100 or so spikes per second common in neurons. Companies like Intel and IBM have developed neuromorphic chips with more than 100,000 neuron-like processors operating in parallel. IBM's TrueNorth chip has one million neurons and is capable of 46 billion synaptic operations per second while using only 70 milliwatts (Akopyan et al. 2015). Intel's Loihi chip has 128 neuromorphic cores each with 1,024 neural units (Davies, 2018). Schuman et al. (2017) survey neuromorphic computing.

Besides speed, the major motivation for neuromorphic computers is power consumption. Whereas a traditional supercomputer can require thousands of watts of power, the brain uses only around 20 watts. The efficiency of computing can be measured by considering the ratio of operations per energy, for example, inferences per joule, neural spikes per joule, or (in supercomputers) floating point operations per second per watt. More abstractly, information efficiency can be measured in bits per joules.

Early experimental results find that neuromorphic computers are indeed more energy efficient than digital computers (Martí, Rigotti, Seok, and Fusi 2016; Blouw, Choo, Hunsberger, and Eliasmith 2019). There are several reasons for such advantages. Neuromorphic machines do much of their computations locally with artificial neurons connected to ones that are physically close to them, requiring much less long-distance transmission of information over wires. Algorithms for learning and memory operate locally so that data do not need to be widely broadcast. Processing elements in neuromorphic computers only emit spikes when required, whereas CPUs are constantly active. Parallelism reduces the number of steps to transform an input into an output. Finally, distributing computation across many processors reduces heat generation so that less energy is required for cooling. Hence, neuromorphic computing is attractive for applications such as biomedical devices where low power consumption and real time operation are desirable.

There are also time-related reasons why neural computers can be more efficient than traditional ones. Neural spiking may not operate as rapidly as traditional silicon chips, but it can operate in synchrony with inputs from processes in the world without any external clock. Time represents itself because of the correspondence between changes in the world and changes in the brain. Neural operations occur in step with changes in the world on the same time scale, although most current neuromorphic chips do not exploit this feature.

Neuromorphic computing illustrates the same energy-information tradeoffs that occur in biology. In general, the more syntactic and semantic information processed by a computer, the more energy it uses. But neuromorphic computers require less energy to perform the same task, so their information/energy ratio is better. As with biology, information processing depends on energy, and energy depends on material substrates.

4.2 Machine learning

Artificial intelligence has made major progress since 2012 because of powerful applications of a neural network technique called deep learning, which has led to spectacularly effective applications such as face and speech recognition (LeCun, Bengio, and Hinton 2015; Goodfellow, Bengio, and Courville 2016). The AI company DeepMind has combined deep learning with reinforcement learning to produce amazing successes such as the world's best Go player (Silver et al. 2018). DeepMind researchers can afford to have their Go player improve by playing itself millions of times, but the resulting information is gained at the expense of a large amount of electrical energy. Schwartz, Dodge, Smith, and Etzioni (2019) describe how the computations required for deep learning have been doubling every few months. Strubell, Ganesh, and McCallum (2019) report the financial and environmental costs of training neural networks for natural language processing.

The rapidly increasing use of electricity for deep learning and other intensive kinds of computing such as cryptocurrency is concerning for reasons that go beyond corporate expense. Standard ways of generating electricity, such as gas power plants, produce large amounts of greenhouse gases that contribute to global warming which threatens to change the Earth's climate in disastrous and irreversible ways in a matter of decades. Reducing energy use for deep learning and other AI applications should therefore be a general social concern, and research is investigating generally how to make computation more energy efficient (Demaine, Lynch, Mirano, and Tyagi 2016).

In contrast to current machine learning, human brains learn from much smaller numbers of examples with energy-efficient processes that make local changes in the strength of synaptic connections. Brains have ways of making learning more efficient by considering relevant subsets of information in data such as ones representing causal relations that circumscribe what might be relevant to human goals. Deep learning has no understanding of causality, so it needs huge numbers of training examples to extract signal from noise. Thagard (2021) examines what it would take to give computers a human-level capacity for causal reasoning.

Attempts are underway to make deep learning more efficient by performing it on neuromorphic computers (Esser et al. 2016). An alternative research strategy is to develop new algorithms that allow AI systems to learn from small numbers of examples using human methods such as abductive inference to causes, analogy, and taking into account background information about variability (Holland, Holyoak, Nisbett, and Thagard 1986). Applications of machine learning have been able to rely on abundant sources of electricity, but constraints such as global warming and portability suggest that AI researchers need to pay attention to energy/information tradeoffs. So deep learning developments point to further connections among inference, energy, and material substrates. Another area of current research concerns the prospects of quantum computing for increasing energy efficiency (Ikonen, Salmilehto, and Möttönen 2017).

So deep learning and neuromorphic computing both show the relevance of energy to understanding and evaluating information processing in computers. As for biological systems, efficient use of energy requires attention to material substrates.

5. Arguments for substrate independence

Substrate independence is often just assumed, but arguments have been used to support its plausibility based on the Church-Turing Thesis and systematic replacement of components. These arguments fail because they neglect to consider energy.

5.1 Church-Turing thesis

In the 1930s, Alonzo Church and Alan Turing independently produced mathematical characterizations of effective computation. Church's lambda calculus and Turing's abstract machine turned out to be equivalent to each other and to Kurt Gödel's characterization of recursive functions. This convergence is taken to support the plausibility of the Church-Turing thesis according to which every effective computation can be carried out by a Turing machine (Copeland 2017). Tegmark (2017, 65) says, "This fact that exactly the same computation can be performed on *any* universal computer means that *computation is substrate independent.*"

The Church-Turing thesis might be taken to provide support for substrate independence because it shows that computation does not depend on a particular kind of computing method. But all three mathematical formulations (Church, Turing, Gödel) ignore time, space, and energy, so the Church-Turing thesis says little about real-world computational devices. Two computers might be equivalent in what they compute but differ markedly in how long they take to do so and how much energy they consume, as we saw in the difference between traditional digital computers and neuromorphic computers. In real life, bits are not free. The mathematical plausibility of the Church-Turing thesis provides no support for the doctrine of multiple realizability, because it does not address the question of how computation can be realized in real physical systems bounded by time and space as well as energy.

5.2 Gradual replacement

Another argument for substrate independence was used by David Chalmers to support the feasibility of uploading conscious minds into computers:

> Here we upload different components of the brain one at a time, over time. This might involve gradual replacement of entire brain areas with computational circuits, or it might involve uploading neurons one at a time. The components might be replaced with silicon circuits in their original location, or with processes in a computer connected by some sort of transmission to a brain. It might take place over months or years, or over hours. If a gradual uploading process is executed correctly, each new component will perfectly emulate the component it replaces, and will interact with both biological and nonbiological components around it in just the same way that the previous component did. So the system will behave in exactly the same way that it would have without the uploading. In fact, if we assume that the system cannot see or hear the uploading, then the system need not notice that any uploading has taken place. Assuming that the original system said that it was conscious, so will the partially uploaded system. (Chalmers 2010, 45–46)

If the parts of one system can be gradually replaced with very different parts without loss of function, then the actual composition or substrate does not matter. Like philosophical thought experiments in general, the brain replacement exercise underspecifies the scientific facts in ways contrived to reach a desired conclusion (Thagard 2014, 2019b). Parts such as neurons do not operate in isolation, but rather in mechanisms where connected parts interact to produce regular changes. Whether the different kinds of parts perform the same function depends not just on the parts but also on the connections and interactions, both of which require energy. If the energy uses of the replacing combination of parts are very different from the energy requirements of the original combination, then it is unlikely that the gradual replacement will work.

Consider the mediocre history of artificial heart replacement. The Jarvik mechanical heart was first implanted in a human in 1982, and many variants have since been produced. Decades later, however, artificial hearts are only used as temporary replacements because of problems that include immune system rejection, durability, and power supply (McKellar 2018).

Similarly, gradual replacement of groups of neurons with silicon chips has substantial obstacles of space, time, and energy. Silicon chips are much larger than neural groups, so they would have difficulty connecting with them within the confines of the skull. Timing is very different in neurons and chips, even neuromorphic ones, so there are major difficulties in coordinating the interactions of chips and neurons. Finally, the problems of systematically replacing the biochemical energy sources of neurons with the electrical energy sources of silicon chips are large. In accord with the five characteristics I identified, use of energy has to satisfy constraints concerning limited supplies, costs of operation, consumption rates, efficiency, and tipping points. All of these are affected by the materials that use energy in different ways.

The gradual replacement story is appealing to those who already believe that substrates are not important to thinking and computation. But attention to the mechanisms that operate in different systems show that actual replacement is problematic because of constraints of space, time, and energy. So the abstract possibility of replacement does not support the plausibility of substrate independence.

The replacement argument encourages the hypothesis of multiple realizability that is often used to show that mental properties cannot be reduced to brain properties. Critics of the idea of multiple realizability have pointed out that cognitive neuroscience has in fact made substantial progress by connecting aspects of human cognition such as vision with neural mechanisms (Bechtel and Mundale 1999; McCauley and Bechtel 2001; Thagard 1986, 2019a). In computer science, it has become increasingly evident that hardware and software are not independent, and that practical success requires that algorithms match hardware. The Church-Turing thesis and replacement argument fail to support substrate independence because of mounting evidence that information and energy depend on material details.

6. Multiple realization and reduction

Place (1956) and Smart (1959) proposed that mental states are identical to brain states, but their proposal was challenged by highly influential arguments from Putnam (1975) and Fodor (1975). Psychology cannot be reduced to neuroscience because mental states such as pain can be realized in different ways, for example, by human brains, non-human brains, and computers. Polger and Shapiro (2016, 2018) provide a thorough critique of this anti-reductionist argument from multiple realization. I largely agree with this critique and will support it with considerations about energy.

According to Polger and Shapiro (2016, 62), multiple realization occurs only if two systems perform relevantly the same function in relevantly different ways. They argue that assumed cases of multiple realization such as vision in humans and octopuses do not actually qualify as multiple realization and therefore do not count against mind-brain identity. My examples from biology and computation show that energy should be taken into account in the evaluation of the relevant similarities and differences. The significance of energy is apparent in considerations of endothermy, information processing, and evidence for mind-body identity.

6.1 Endothermy

Polger and Shapiro (2016, 142) mention that endothermic (warm-blooded) animals must acquire large amounts of energy and therefore consume much more food than ectothermic (cold-blooded) organisms. There are indeed important relationships among metabolism, body-brain size ratios, and neuronal densities (Yu, Karbowski, Sachdev, and Feng 2014). Larger brains with more neurons consume more energy, which is a challenge for organisms that need energy to keep themselves warm. Accordingly, ectothermic animals in tropical climates have larger brains than ectothermic animals in colder climates. A key factor in the development of large-brained endotherms is the rapid increase in the number of glial cells that are less energy-demanding than neurons.

These findings show that comparisons of psychological functions such as vision across different organisms cannot ignore energy considerations. Such functions operate in environments that include factors of external and internal temperature that interact with the neural mechanisms that carry out the psychological functions. Hence, determination of whether psychological functions are multiply realized needs to include energy assessments. Mammal and bird brains operate in warm-blooded bodies unlike the cold-blooded bodies of molluscs, reptiles, and fish, which can affect basic neural operations such as the balance between number of synaptic connections per neuron and strength of weights between neurons (Yuan et al. 2018).

6.2 Information processing

Polger and Shapiro (2016, ch. 8) provide an extensive discussion of information-processing models in cognitive science. They recognize that such models have been a major motivation for functionalist accounts in the philosophy of mind: if human thinking is information processing, which can also be done by computers, then mental functions are multiply realized.

But what is information? Recent surveys in the Stanford *Encyclopedia of Philosophy* provide no general answer (Adriaans 2018; Floridi 2015; Godfrey-Smith and Sterelny 2016). Dictionary definitions of information as facts or knowledge are not informative. Floridi 2010 defines information as data that are well-formed and meaningful, but this is of limited use when dictionaries commonly define a datum as a piece of information and Floridi's own account of a datum as something distinct is vague. The elegant probabilistic characterization of information by Shannon and Weaver (1949) has little application to biological systems.

A better way of characterizing information is to specify the classes of mechanisms by which it operates, which include representation, collection, storage, retrieval, evaluation, transformation, sending, and receiving. All of these operate on different carriers of information that include utterances, sentences, gestures, pictures, patterns of neural firing, computational processes, and genetic structures made of DNA. I will illustrate the eight mechanisms using neural patterns.

Representation allows carriers to stand for situations in the world. For example, different groups of neurons are tuned, either innately or by learning, to different aspects of the world such as colors. Such tuning enables collection of new information by interaction with the world, for example, when the retina sends signals from the eye to the brain where storage takes place by generation of synaptic connections. For ongoing use, the information must be retrieved from storage, which the brain accomplishes by firing the neurons based on their synaptic connections. Whether a carrier of information is collected, stored, and retrieved depends on its relevance to an organism's goals such as survival and reproduction, where relevance is assessed by an evaluation mechanism. In brains, evaluation is largely performed by emotional brain areas such as the amygdala, nucleus accumbens, and ventromedial prefrontal cortex. Information processing also includes making inferences that transform carriers into new and useful ones using mechanisms that range from learning by association to analogical reasoning. For many species of organisms such as humans, information is communicated between individuals, which requires mechanisms for both sending and receiving different kinds of carriers. Humans have neural processes for producing utterances as well as for interpreting the utterances of others.

All these mechanisms require energy to keep the parts working, to keep the parts connected, and especially to accomplish the interactions of parts that produce regular changes. Neurons need energy to survive, maintain synaptic connections with other neurons, and interact by processes of excitation and inhibition. Hence, determination of whether information processing is multiply realized must consider the extent to which energy constitutes a relevant similarity or difference. Because neurons and computer chips have mechanisms that use energy in very different ways, we cannot assume that abstract information processing supports functionalism.

Information processing requires attention to speed as well as inputs and outputs. If the eight information mechanisms operate too slowly, then an organism will not be able to accomplish its mental functions in its environmental context. For example, if an animal is too slow in its transformations of neural patterns, then it is not able to recognize food and predators fast enough to survive and reproduce, with resulting extinction. Fast processing is heavily dependent on energy in the same way that running consumes energy more rapidly than walking. Hence, debates about the multiple realization of information processing should take into account time/energy tradeoffs.

6.3 Identity theory

Polger and Shapiro's critique of multiple realization arguments is in defense of their version of a mind-body identity theory, but Endicott (2017) remarks that they do not advance specific mind-body identity claims. In contrast, Thagard (2019a, b) specifies neural mechanisms for a host of important mental processes that include perception, imagery, concepts, rules, analogies, emotions, consciousness, actions, intentions, language, creativity, and the self. The requisite mechanisms are based on Eliasmith's (2013) Semantic Pointer Architecture, which pays close attention to the brain's energy usage. Hence arguments for mind-brain connections can in part be energy-based.

At first glance, it might seem that the Semantic Pointer Architecture supports multiple realization because its large-scale simulations with millions of artificial neurons have largely been run on conventional computers (Eliasmith et al. 2012). But Eliasmith and Stewart (2020) explain why the most recent versions are being run on neuromorphic chips that mimic the energy efficiency of the brain. They estimate that an artificial neural network the size of a human brain would take half a gigawatt of power using today's computers, equivalent to the power produced by a nuclear power plant. Use of Intel's Loihi neuromorphic chip is far more energy efficient and makes increasingly large simulations feasible.

Polger and Shapiro (2016, 158) remark that connectionist neural networks can be built out of most anything, but considerations of energy and time show that multiple realization fails for them as well as for many other proposed systems. A neural network built out of toilet paper would be too slow, fragile, and energy demanding to be of any use. Future defenses of the richness of mind-brain connections can take into account ways in which the effectiveness of psychological functions depends on the energy-efficiency of the brain and neuromorphic computer chips. The use of specialized machines for brain simulations does not support multiple realization because there are still important energy differences between brains and neuromorphic computers, for example, in the role of glial cells and different kinds of neurotransmitters.

I have shown how energy considerations concerned with endothermy, information processing, and mind-brain identities support the rejection of multiple realization arguments for functionalism. The philosophical debate is ongoing, as serious responses to Polger and Shapiro (216) have been given by Aizawa (2017), Chirimuuta (2018), and others. Garson (2003) and Piccinini (2015) maintain that information and computation are "medium independent", another term for substrate independent. I hope that future discussions will include close attention to energy requirements.

7. Conclusion

According to Tegmark (2017), "matter doesn't matter" to computation, intelligence, and consciousness. In contrast, I have argued that attention to the energy requirements of different kinds of mechanisms shows that the constitution and operation of both biological and computational systems are highly dependent on energy. Any information processing system requires energy for its sensory inputs, internal processing, and outputs by communication and action. Acquisition of energy also requires energy, which therefore must be used efficiently if an organism or computer is to be effective enough to survive. In addition to the energy required for the interactions between parts in a mechanism, energy is required for the forces and processes that keep parts connected (Findlay and Thagard 2012).

Substrate independence is an empirical claim poorly supported by arguments such as the Church-Turing thesis and the gradual replacement thought experiment. Energy considerations in biology and computation make general substrate independence implausible because different forms of matter have markedly different costs as measured by ratios of information processing operations to energy consumed. There may, however, be special cases where transformation between substrates does not violate energy constraints, but my biological and computational examples show that such transformations cannot be taken for granted.

These examples provide substantial support for the two premises of my general argument:

1. Real world information processing depends on energy.
2. Energy depends on material substrates.

Here "A depends on B" means "changes in B cause important changes in A". I have shown how energy differences cause important differences in information processing: the contribution of metabolism to perception in living systems, the effects of energy availability on brain evolution, the speed and power consumption advantages of low-energy neuromorphic computers, energy challenges of machine learning, and the role of energy in information mechanisms. In all of these cases, different material substrates have causal effects on energy usage and hence on information processing. So, information processing depends on material substrates.

The resulting implausibility of substrate independence has important consequences for current debates. The energy dependence of thinking and computation does not imply that artificial intelligence is impossible, but it does show that we cannot automatically assume that human thought processes can be converted into computational mechanisms. Whether such conversion can be realized is an empirical question that can only be answered by ongoing investigation. Particularly problematic are aspects of thinking that are closely tied to the biological mechanisms of human bodies, including perception, emotion, and consciousness (Barsalou 2008; Pepperell 2018).

Because computers are unlikely ever to emulate the energy operations of human bodies, AI will have to find ways of duplicating or working around the major contributions that imagery, emotion, and conscious experience make to human thought. Because these functions are also important to the ethical lives of humans, abandonment of substrate independence casts doubt on whether computer systems will ever deserve the same moral status as humans (Thagard 2021).

If substrate independence is false, then the prospects for uploading minds to computers become much less appealing. Merely converting neurons and synapses to their computational equivalents will not replicate the neural mechanisms that depend on biochemical, glucose-based energy. Neuroscience does not yet have an account of how neural operations generate consciousness, although there are suggestions about relevant mechanisms, including information integration (Tononi, Boly, Massimini, and Koch 2016), broadcasting to a neural blackboard (Dehaene 2014), and competition among neural representations called semantic pointers (Thagard and Stewart 2014). These mechanisms can be combined in a common model (Thagard in press). Each of these mechanisms is energy intensive and therefore might depend on the particular kind of energy used by brains. Without the dubious assumption of substrate independence, it is mere faith that uploading brains to computers will maintain human-like consciousness.

Challenging substrate independence undercuts speculations that we are currently living in a computer simulation. Some future computer simulations might turn out to be conscious, but the very different mechanisms and energy operations of computers and brains make it equally plausible that simulations of thinking will not achieve consciousness, just as simulations of ocean waves are not wet.

Finally, the undermining of claims about substrate independence and multiple realization has implications for understanding the relation between psychology and neuroscience. Rejection of substrate independence does not imply reversion to dualist claims that mind and thinking require non-material substances. Viable alternatives to dualism include mind-body identity, the claim that all human mental processes are neural mechanisms. The important interconnections among information, energy, and material substrates show that mind-body identity is not refuted by arguments about multiple realization. A more complex but still materialist alternative is that mind results from multilevel interacting mechanisms that are molecular and social as well as neural (Thagard 2019a, b). Either way, taking energy seriously in both biological and computational systems points to a deeper understanding of minds, bodies, and computers.

References

Adriaans, Pieter. 2018. "Information." *Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/entries/information/

Aizawa, Kenneth. 2018. "Multiple Realization and Multiple "Ways" of Realization: A Progress Report." *Studies in the History and Philosophy of Science* 68:3–9.

Akopyan, Filipp, Jun Sawada, Andrew Cassidy, Rodrigo Alvarez-Icaza, John Arthur, Paul Merolla, Nabil Imam, Yutaka Nakamura, Pallab Datta, and Gi-Joon Nam. 2015. "Truenorth: Design and Tool Flow of a 65 mw 1 Million Neuron Programmable Beurosynaptic Chi." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 34 (10):1537–57.

Barsalou, Lawrence W. 2008. "Grounded Cognition." *Annual Review of Psychology* 59:617–45.

Bechtel, William. 2008. *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. New York: Routledge.

Bechtel, William, and Jennifer Mundale. 1999. "Multiple Realizability Revisited: Linking Cognitive and Neural States." *Philosophy of Science* 66:175–207.

Bickle, John. 2020. "Multiple Realizability." *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/multiple-realizability/

Blackford, Russell, and Damien Broderick. 2014. *Intelligence Unbound: The Future of Uploaded and Machine Minds*. Chichester: John Wiley & Sons.

Blouw, Peter, Xuan Choo, Eric Hunsberger, and Chris Eliasmith. 2019. "Benchmarking Keyword Spotting Efficiency on Neuromorphic Hardware." *Proceedings of the 7th Annual Neuro-inspired Computational Elements Workshop.*

Bostrom, Nick. 2003. "Are We Living in a Computer Simulation?" *The Philosophical Quarterly* 53:243–55.

Bostrom, Nick, and Eliezer Yudkowsky. 2014. "The Ethics of Artificial Intelligence." In *The Cambridge Handbook of Artificial Intelligence*, ed. Keith. Frankish and William M. Ramsey, 316–34.

Chalmers, David J. 2010. "The Singularity: A Philosophical Analysis." *Journal of Consciousness Studies* 17:7–65.

Chirimuuta, Mazviita. 2018. "Marr, Mayr, and MR: What Functionalism Should Now Be About." *Philosophical Psychology* 31 (3):403–18.

Coolidge, Frederick L., and Thomas Wynn. 2018. *The Rise of Homo Sapiens.* 2nd ed. Oxford: Oxford University Press.

Copeland, Brian Jack. 2017. The Church-Turing thesis. *Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/entries/church-turing/

Corballis, Michael C. 2011. *The Recursive Mind: The Origins of Human Language, Thought, and Civilization.* Princeton: Princeton University Press.

Craver, Carl F., and Lindley Darden. 2013. *In Search of Mechanisms: Discoveries Across the Life Sciences.* Chicago: University of Chicago Press.

Davies, Mike, Narayan Srinivasa, Tsung-Han Lin, Gautham Chinya, Yongqiang Cao, Sri Harsha Choday, Georgios Dimou, Prasad Joshi, Nabil Imam, and Shweta Jain. 2018. "Loihi: A Neuromorphic Manycore Processor with On-Chip Learning." *IEEE Micro* 38 (1):82–99.

Dehaene, Stanislas. 2014. *Consciousness and the Brain: Deciphering How the Brain Codes our Thoughts.* New York: Viking.

Demaine, Erik D., Jayson Lynch, Geronimo J. Mirano, and Nirvan Tyagi. 2016. "Energy-Efficient Algorithms." *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science.*

Eliasmith, Chris. 2013. *How to Build a Brain: A Neural Architecture for Biological Cognition.* Oxford: Oxford University Press.

Eliasmith, Chris, Terrence C. Stewart, Xuan Choo, Trevor Bekolay, Travis DeWolf, Yichuan Tang, and Daniel Rasmussen. 2012. "A Large-Scale Model of the Functioning Brain." *Science* 338:1202–5.

Eliasmith, Chris, and Terry Stewart. 2020. "Scaling, Low Power Needed for a Neurmorphic Future." *EE/Times.* Retrieved from https://www.eetimes.com/scaling-low-power-needed-for-a-neuromorphic-future/#

Endicott, Ronald. 2017. "The Counter-Revolution over Multiple Realization." *Metascience* 26 (2):229–32.

Esser, Steven K., Paul A. Merolla, John V. Arthur, Andrew S. Cassidy, Rathinakumar Appuswamy, Alexander Andreopoulos, David J. Berg, Jeffrey L. McKinstry, Timothy Melano, Davis R. Barch, Carmelo di Nolfo, Pallab Datta, Arnon Amir, Brian Taba, Myron D. Flickner, and Dharmendra S. Modha. 2016. "Convolutional Networks for Fast, Energy-Efficient Neuromorphic Computing." Proceedings of the National Academy of Sciences 113 (41):11441–6.

Findlay, Scott D., and Paul Thagard. 2012. How Parts Make up Wholes. *Frontiers in Physiology* 3. https://www.frontiersin.org/articles/10.3389/fphys.2012.00455/full

Floridi, Luciano. 2010. *Information: A Very Short Introduction.* Oxford: Oxford University Press.

Floridi, Luciano. 2015. "Semantic Conceptions of Information." *Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/entries/information-semantic/

Fodor, Jerry. 1975. *The Language of Thought.* New York: Crowell.

Garson, Justin. 2003. "The Introduction of Information into Neurobiology." *Philosophy of Science* 70 (5):926–36.

Glennan, Stuart. 2017. *The New Mechanical Philosophy*. Oxford: Oxford University Press.

Godfrey-Smith, Peter, and Kim Sterelny. 2016. Biological Information. *Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/entries/information-biological/

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. 2016. *Deep Learning.* Cambridge, MA: MIT Press

Greene, Preston. 2018. "The Termination Risks of Simulation Science." *Erkenntnis* 85 (2):489–509.

Herculano-Houzel, Suzana. 2016. *The Human Advantage: A New Understanding of How our Brain Became Remarkable*. Cambridge, MA: MIT Press.

Holland, John H., Keith J. Holyoak, Richard E. Nisbett, and Paul R. Thagard. 1986. *Induction: Processes of Inference, Learning, and Discovery*. Cambridge, MA: MIT Press.

Ikonen, Joni, Juha Salmilehto, and Mikko Möttönen. 2017. "Energy-Efficient Quantum Computing." *Quantum Information* 3 (1). https://doi.org/10.1038/s41534-017-0015-5.

Kurzweil, Ray. 2005. *The Singularity Is Near: When Humans Transcend Biology*. New York: Viking.

Kurzweil, Ray. 2012. *How to Create a Mind*. New York: Viking.

Lane, Nick, and William Martin. 2010. "The Energetics of Genome Complexity." *Nature* 467 (7318):929–34.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. "Deep Learning." *Nature* 521 (7553):436–44.

Levin, J. 2018. Functionalism. *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/functionalism/

Martí, D., M. Rigotti, M. Seok, and S. Fusi. 2016. "Energy-Efficient Neuromorphic Classifiers." *Neural Computing* 28 (10):2011–44.

McCauley, Robert N., and William Bechtel. 2001. "Explanatory Pluralism and the Heuristic Identity Theory." *Theory & Psychology* 11:736–60.

McKellar, Shelley. 2018. *Artificial Hearts: The Allure and Ambivalence of a Controversial Medical Technology*. Balltimore: JHU Press.

Mead, Carver. 1990. "Neuromorphic Electronic Systems." *Proceedings of the IEEE* 78 (10):1629–36.

O'Connor, Mary I., Matthew W. Pennell, Florian Altermatt, Blake Matthews, Carlos J. Melián, and Andrew Gonzalez. 2019. "Principles of Ecology Revisited: Integrating Information and Ecological Theories for a More Unified Science. *Frontiers in Ecology and Evolution*. https://www.frontiersin.org/articles/10.3389/fevo.2019.00219/full

Pepperell, Robert. 2018. "Consciousness as a Physical Process Caused by the Oganization of Energy in the Brain." *Frontiers in Psychology* 9:2091. https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02091/full

Piccinini, Gualtiero. 2015. *Physical Computation: A Mechanistic Account*. Oxford: Oxford University Press.

Place, Ullin T. 1956. "Is Consciousness a Brain Process?" *British Journal of Psychology* 47:44–50.

Polger, Thomas W., and Lawrence A. Shapiro. 2016. *The Multiple Realization Book*. Oxford University Press.

Polger, Thomas W., and Lawrence A. Shapiro. 2018. "Responses to Critics." *Philosophical Psychology* 31 (3):446–57.

Putnam, Hilary. 1975. *Mind, Language, and Reality*. Cambridge: Cambridge University Press.

Raichle, Marcus E., and Debra A. Gusnard. 2002. "Appraising the Brain's Energy Budget." *Procedings of the National Academy of Sciences* 99 (16):10237–9.

Schneider, Susan. 2019. *Artificial you: AI and the future of your mind*. Princeton, NJ: Princeton University Press.

Schuman, Catherine D., Thomas E. Potok, Robert M. Patton, J. Douglas Birdwell, Mark E. Dean, Garrett S. Rose, and James S. Plank. 2017. "A Survey of Neuromorphic Computing and Neural Networks in Hardware." *arXiv preprint arXiv:1705.*06963.

Schwartz, Roy, Jesse Dodge, Noah A. Smith, and Oren Etzioni. 2019. "Green AI." *arXiv preprint arXiv:1907.10597.*

Shannon, Claude, and Warren Weaver. 1949. *The Mathematical Theory of Communication.* Urbana: University of Illinois Press.

Silver, David, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, and Thore Graepel. 2018. "A General Reinforcement Learning Algorithm that Masters Chess, Shogi, and Go through Self-play." *Science* 362 (6419):1140–4.

Smart, John J. C. 1959. "Sensations and Brain Processes." *Philosophical Review* 68:141–56.

Strubell, Emma, Ananya Ganesh, and Andrew McCallum. 2019. "Energy and Policy Considerations for Deep Learning in NLP." *arXiv preprint arXiv:1906.02243.*

Szathmáry, Eörs. 2015. "Toward Major Evolutionary Transitions Theory 2.0." Proceedings of the Natlonal Academy of Sciences 112 (33):10104–11.

Tegmark, Max. 2017. *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York: Alfred A. Knopf.

Thagard, Paul. 1986. "Parallel Computation and the Mind-Body Problem." *Cognitive Science* 10:301–18.

Thagard, Paul. 2014. "Thought Experiments Considered Harmful." *Perspectives on Science* 22:288–305.

Thagard, Paul. 2019a. *Natural Philosophy: From Social Brains to Knowledge, Reality, Morality, and Beauty*. New York: Oxford University Press.

Thagard, Paul. 2019b. *Brain-mind: From Neurons to Consciousness and Creativity*. New York: Oxford University Press.

Thagard, Paul. 2021. *Bots & Beasts: What Makes Machines, Animals, and People Smart?* Cambridge, MA: MIT Press.

Thagard, Paul. In press. *Balance: How It Works and What It Means*. New York: Columbia University Press.

Thagard, Paul, and Terrence C. Stewart. 2014. “Two Theories of Consciousness: Semantic Pointer Competition vs. Information Integration.” *Consciousness and Cognition* 30:73–90.

Tononi, Giulio, Melanie Boly, Marcello Massimini, and Christof Koch. 2016. “Integrated Information Theory: From Consciousness to its Physical Substrate.” *Nature Reviews Neuroscience* 17 (7):450–61.”

Wrangham, Richard W. 2009. *Catching Fire: How Cooking Made us Human*. New York: Basic Books.

Yu, Yuguo, Jan Karbowski, Robert NS Sachdev, and Jianfeng Feng. 2014. “Effect of Temperature and Glia in Brain Size Enlargement and Origin of Allometric Body-Brain Size Scaling in Vertebrates.” *BMC Evolutionary Biology* 14 (1):1–14.

Yuan, Ye, H. Huo, Hong Zhao, Peng Liu, Jian Liu, Jaxing, Xing, Fu, and Tao Fang. 2018. “Constraints of Metabolic Eenergy on the Number of Synaptic Connections of Neurons and the Density of Neuronal Networks.” *Frontiers in Computational Neuroscience* 12. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6256250/

Cite this article: Thagard, Paul. 2022. “Energy Requirements Undermine Substrate Independence and Mind-Body Functionalism.” *Philosophy of Science* 89 (1):70–88. https://doi.org/10.1017/psa.2021.15