



There is no “I” in “AI”

Ashkan Farhadi^{1,2}

Received: 31 August 2020 / Accepted: 12 December 2020 / Published online: 15 January 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd. part of Springer Nature 2021

Abstract

With recent advancements in technology and computer science, we have reached a point where we can clearly state that thinking is no longer the exclusive privilege of living minds. Artificial intelligence (AI) can gather and process information in a manner fairly similar or even superior to our thinking process. AI can use this processed information in a reasoning process to make decisions and execute them. However, what makes our mind distinct from AI is the addition of “I,” that is, an entity composed of two mental processes of awareness-based choice selection (ABCS) and discretionary selection of information for attention (DSIA) that are the fundamental core function of decision making and awareness processes, respectively. ABCS allows the coexistence of free will alongside—but unrestrained by—the chain of causality in the decision-making process with the help of awareness. “I” also allows for DSIA, which is a key step in the awareness process. This intertwined process is crucial for our consciousness. Similar to AI, the mind is devoid of any form of self-recognition without awareness. Self-awareness is a collaborative function of “I” and the mind. “I” is instrumental in the sense of self-awareness, but on its own, it is selfless. Additionally, “I” is the heart of the decision-making process. Therefore, AI is missing “I,” a selfless master of the mind.

Keywords Artificial intelligence · Mind · Consciousness · Awareness · Free will · Self-awareness

1 Introduction

We live in a deterministic world. As the light of science shines brighter over subjects and phenomena, the shadow of the unknown and unexplainable shrinks. In the light of science, we see how events and phenomenon are shackled to the chain of cause and effect and this deterministic relationship makes us capable of predicting the outcomes of events with increasing precision. To explore the shadows that are yet to see the light of deterministic science, we resort to statistics to make sense of events and allow ourselves to retain our ability to predict the outcomes of events. It seems that, in a deterministic world of science where any event is preceded by another event in a chain of causality, there would be no room for free will (determinism). Some argue that even with a deterministic backdrop, there is still room to exercise free will. By referring to indeterministic properties

proposed by quantum mechanics, some have proposed that not all aspects of our universe follow deterministic rules. Therefore, deterministic choices do not have to end up in a deterministic selection by us (compatibilism). Some have equated this indeterministic selection process of choices to free will. Based on this perspective, we are another cog in the deterministic universe that turns and twists as has already been determined, but sometimes randomly goes through an unpredictable hiccup and fails to produce the expected result, thanks to the indeterministic laws of quantum mechanics. In this scenario, we are nothing but a fancy computer-like creature that follows a very sophisticated algorithm that has been programed in us by years of knowledge, experience, and desires, and the chemical properties of our brain. Our brain is a computer that misfires randomly and produce some unexpected results from time to time. The difference between natural intelligence (NI) and artificial intelligence (AI) is nothing but the amount of stored information, the processing speed of information, and/or the sophistication of the algorithms that process the information. On the other end of spectrum, there are proposals for a metaphysical process in the mind that allows detachment from the deterministic world and provides us with the gift of free will

✉ Ashkan Farhadi
afarhadi@memorialcare.org

¹ University of California, Irvine, USA

² Memorial Care Medical Group, 18035 Brookhurst St #2100,
Fountain Valley, CA 92708, USA

(libertarianism). In this perspective, there would be a sharp line separating NI and AI by considering a metaphysical property for the mind that is absent in AI. Yet, an alternative to these perspectives is the proposal of the trilogy theory of mind that is further explained in this manuscript. However, before we delve deeper into what makes NI different from AI, I want to discuss the background of the advancement of AI.

Since the adaptation of the Turing test in 1950 in his article “Computing and Machinery Intelligence,” discussion has continued about the true meaning of an intelligent behavior and whether a machine could actually think or at least respond in a way that would lead an observer to believe that the machine could think.

Following Turing’s interesting imitation game, a group of scientists gathered in a workshop in Dartmouth in 1956 and concluded that “every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it.” In Turing’s experiment, intelligence was oversimplified as the ability of a machine to use a similar vocabulary as a human in a proper manner. In his prediction, Turing predicted that in 50 years there would be computers with processing capabilities that could fool 70% of interrogators into believing that they were communicating with a live person. His prediction proved inaccurate and no computer was able to pass his test in half a century. But in 2014, in an experiment held at the University of Reading, Eugene Goostman, a computer program simulating a 13-year-old Ukrainian boy, successfully passed the Turing test for the first time.

With further advancements in computer processing abilities, we currently know that the computer can be considered intelligent since so many experiments have confirmed that computer responses in chat rooms are indistinguishable from human responses by other participants. We all have experienced firsthand that when we pose a question to our iPhone, Siri responds in such a smart and sometimes political manner that we cannot help thinking there is an element of thought behind its response.

The question is where we can truly draw a line between the action of a NI and AI. Some scientists proposed that the mere proper or logical response to an action is not good enough and there should be some modification of the response based on knowledge and past experience for it to be considered an intelligent response (Russell and Norvig 2003). We may argue that even a shaving machine could be considered intelligent, since by averaging the shaving time each time, it corrects itself for the number of shavings per its charged time after each use.

Even though our machines at this time are not capable of performing complicated tasks similar to our minds, some scientists have argued that with advancement in processing capabilities and expansion of knowledge and memory

capabilities of future computers, any action performed by a neuron that is ultimately based on physical/chemical properties should be theoretically reproducible by a machine (Dreyfus 1972). Others have argued that machines could never reproduce brain function. The pioneer of this thought was Kurt Gödel, with his “incompleteness theorem” that was proposed in 1931. In his proposal, a machine may be able to do an arithmetic task based on the provided formula but cannot prove whether the formula itself is true. In other words, a system cannot prove its own consistency. This notion was further elaborated by Lucas (1961) and was later adapted by Roger Penrose into an anti-mechanist idea. Based on this idea, mathematical reasoning strength of the brain could not be duplicated with any machine since unlike the brain, a machine cannot prove a theorem that is not known to it. This may in a sense translate into the ability to make new knowledge and add to what we already know. Considering the computing powers of today’s AI, the anti-mechanist idea could not be considered a differentiating point between AI and human intelligence. Modern AI can be creative or able to come up with ways to solve problems for which it was not originally trained, as with the Euroqa computer at Cornell University (Katz 2009).

There is no question that the current AI are capable of performing several processes such as saving and accessing data, analyzing data, expanding on their current information, gathering information from various inputs and executing tasks based on the received information which could resemble known mental functions such as preserving and recalling memories, reasoning, accruing knowledge and experiences, sensing the surrounding environment and reacting properly to the stimuli, respectively, even better than humans. In the modern era of AI, the majority of thought centers on having consciousness as the differentiating point between the NI and AI. In another words, machines can never be aware of their actions. This was nicely elaborated by John Searle, who concluded in his “Chinese room” paradigm that a machine can perfectly translate English to Chinese for someone who does not know the language, but the machine itself does not understand Chinese (Searle 1980). Turing also reached the same conclusion and pointed out that a machine may not become the subject of its own thought (Turing 1950). The caveat to these arguments is the lack of any objective evidence for subjective phenomena such as understanding. In other words, how can we objectively determine if the machine does or does not understand Chinese aside from the responses that we could get from that machine? How could we appreciate someone’s subjective experience aside from self-explanation, verbal cues or particular actions? But couldn’t an AI pretend that it understands Chinese? The same is true regarding self-awareness. Therefore, even though we may think self-awareness is not possible for AI, we cannot prove its existence or the lack of thereof by any

means but asking questions and judging the appropriateness of the responses, even though we know that AI has been able to successfully pass any forms of advanced Turing tests with proper programming. Also, for those who have a purely materialistic view of the brain and see our consciousness as nothing but a property of our physical brain with its neurons and networks, it is just a matter of time till we incorporate all mental functions including consciousness into AI.

The other differentiating point may be the lack of emotions and their subjective counterpart feeling in AI. We all know that current AI lack any forms of emotion or emotional intelligence. We tend to believe that emotion is reserved for humans and some animals. We also know how emotions have the potential to modify our mental functions, reasoning, decision making and behavior and be considered emotional intelligence. However, could we possibly differentiate AI from a human using this parameter? Regarding emotion, we could easily imagine that the objective effects of emotion on the human body, such as sweating, elevated heart rate and dilation of pupils, could be programmed into an AI. Further, a well-programmed AI could make you believe that it has feelings similar to the argument for awareness presented above. On the other hand, emotional intelligence, just like any other forms of information, could be incorporated into the programming of a smart AI. Therefore, differentiating our mind and a well-programmed AI using this parameter would be very difficult, if not impossible.

Another fundamental differentiating factor between AI and NI may reside in the autonomy to make a decision. There is a library's worth of information on the subject of free will. However, this term is fully loaded with many connotations. As I explain later in the manuscript, there are several stages for decision making and to be more specific I will not use the term free will. Instead, I introduce a new term, awareness-based choice selection (ABCS) which is the heart of the decision-making process in NI and stands in sharp contrast to selection of a choice based on an algorithm (SCBA) that is the current foundation of the decision-making process in AI. We all work with many forms of AI that have autopilot or autonomous features. A self-driving car is not merely a concept anymore and there are robots that decide if it is time to charge themselves and find a plug to connect to charge. Could that be considered an example of forms of autonomy for AI? Even though these levels of decision making are a giant leap of progress toward self-autonomy, there is still a large gap for AI to reach to full autonomy (Markoff 2009). AI cannot make a choice without a preplanned matrix of information, or an algorithm. AI can only make decisions based on the predetermined algorithm—SCBA—that was programmed into it. Even if the reasoning process in AI could bring multiple choices to the decision-making platform, only one will be selected based on the SCBA. One may argue that AI is capable of selecting

different choices every time by adding a random selection process to final choice selection. However, this is on its own a predetermined algorithm. This may well be similar to the idea of choice and chance presented by James (1884) as the main method of decision making. Based on this perspective, neither humans nor AI have much discretion beyond predetermined choices and random chances.

Below, I will describe a new perspective called the trilogy theory of mind, or “trilogy” for short. This new theory is an attempt to further characterize mental function by carving out an independent entity from the mind, called “I.” Based on this theory, we are a union of “I,” our mind and our bodies. “I” in the trilogy is not the true self but an intertwined mental process of DSIA and ABCS and serves as the fundamental distinction between NI and AI. “T” in the trilogy is a selfless master of the mind (Fig. 1).

Before describing trilogy, as a literally starting point, I would like to present the story of creation adapted from the Old Testament. After Adam and Eve were created, they were placed in the Garden of Eden. They roamed around Paradise and lived heavenly lives, innocently enjoying all the amenities until Eve was tricked by a serpent and tempted to taste the fruit of the forbidden tree of knowledge. Despite God's explicit orders to not eat from the tree, Eve took a bite from the fruit and gave it to Adam, who also had a bite. Later, when God arrived at the garden, Adam and Eve hid behind some trees. God asked Adam where he was and Adam replied that he heard a voice and hid because he was naked. It did not take God much time to figure out that Adam had eaten from the tree of knowledge. God stated that the man was now like one of us, knowing good and evil. Therefore, punishment was inevitable and God banished them from the Garden of Eden to lead a mortal life on earth as humans.

At first glance, this story is about man committing the sin of disobeying God and being punished and sent down to earth to live as the first human. However, one can see a silver lining in the story. The innocent Adam and Eve are examples of humanoids that were created as a programmed being that were intended to do what they were told throughout their heavenly lives. However, they exerted their free will and made a decision that changed their lives—and ours—forever. They ate the fruit of the tree of knowledge of good and evil. It is unclear whether the fruit gave them more knowledge, but it certainly gave them awareness. After eating the fruit, they became aware of their nakedness. Because of these two revelations—discretion and awareness—Adam and Eve became God-like, while leaving no option for God but to exile them from Paradise into earth as human.

It seems what turned a humanoid angel into a human was nothing but awareness and discretionary decision-making which are two pillars of the entity called “I” in trilogy. In this theory, “I” sits at the border of the conscious and unconscious minds as well as the realm of subjectivity and

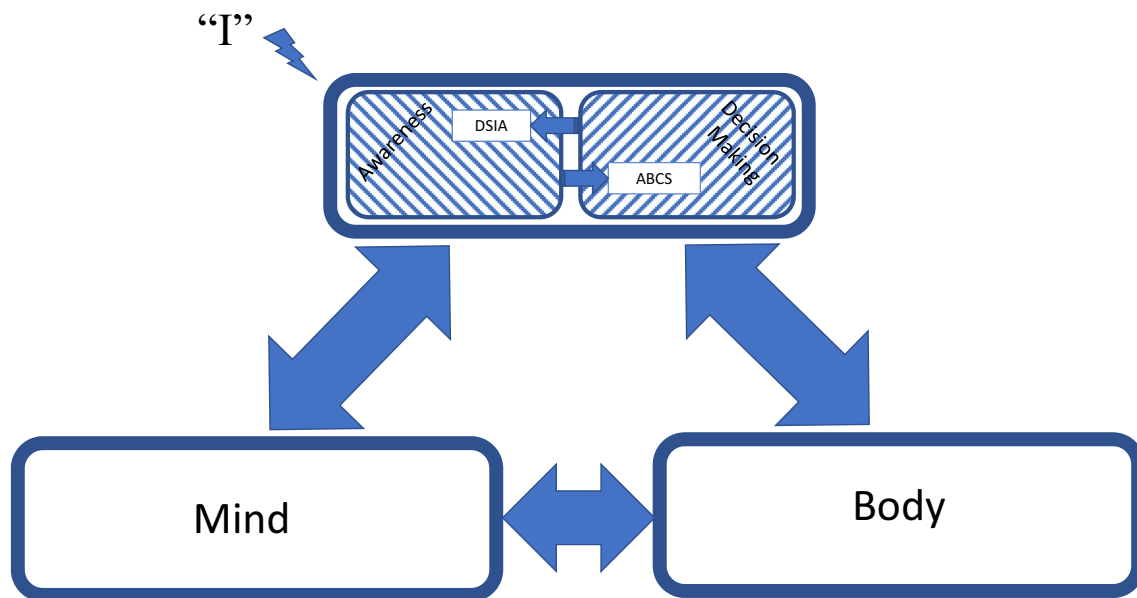


Fig. 1. Based on trilogy we are a union of “I,” our mind and our bodies. “I” composed of two mental process, discretionary selection of information for attention (DSIA) and awareness based choice selec-

tion (ABCS) that are the core of awareness and decision making processes, respectively

objectivity. “I” is instrumental in transformation of *objective information* into *subjective experience* in the process of awareness, and awareness makes ABCS possible. Neither awareness nor decision making can occur independently of one another, and the result of this intertwined process of ABCS and DSIA in “I” is the essence of consciousness, which separates us from AI.

2 There is no “I” in AI

Based on trilogy, the difference between NI and AI is not in a quantitative magnitude of processing power or stored information that one assumes will eventually be overcome with more advanced computational capability or the memory capacity of modern computers. The main difference is in the lack of the “I” in AI that furnishes NI with an intertwined process of DSIA and ABCS. To elaborate further on this topic and explain the differences between AI and our mind in detail, we examine some aspects of awareness and decision-making.

2.1 Decision-making process

There are many similarities and dissimilarities in the decision-making process of our minds and AI (Fig. 2). To point out these differences and similarities, we need to detail the different stages of the decision-making process. The stages of the decision-making process span over our mind and “I.” It should be pointed out that the stages that are proposed

here are based on introspection which can only reveal information regarding decision-making process from a subjective perspective and therefore, cannot depict an empirical perspective of the decision-making process.

2.1.1 Preselection stage

This stage is processed in our mind. For any decision to be made, we gather preliminary information. The information may be presented raw or it may be processed and turned into intelligence. The information could be received from many sources, such as sensory inputs, motor information, memories, knowledge, beliefs, morals, virtues, values, desires, and the thinking process, and could be collectively called informative intelligence. Up to this point we generally share this process with AI. However, there is another category of intelligence that AI does not use in decision making, and that would be emotional intelligence, including our moods, emotions, physical and physiological status (pain, tiredness, hunger, physiological urges), and self-esteem. After our minds gather all the intelligence, the processed information goes through an analytic process called reasoning to create a matrix of information. During the reasoning process, we assign weight or value to our options, and this results in generating the best option among all options. However, the counter-reasoning process is usually a parallel but separate process from reasoning that generates the second or third best option. If the decision-making entity is the same as the entity that does the reasoning (such as AI) there would be no need for counter-reasoning. However, if the choice

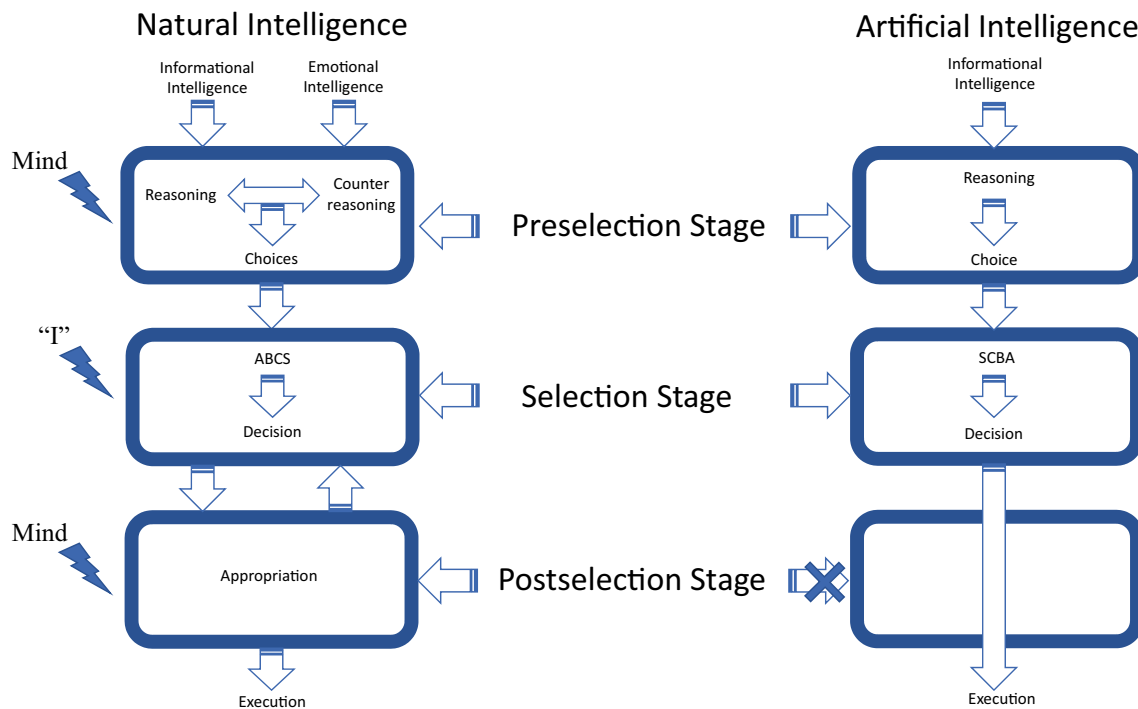


Fig. 2. Similarities and differences of the decision-making process in natural intelligence (NI) versus artificial intelligence (AI). In NI we have awareness-based choice selection (ABCS) while in AI we have

selection of a choice based on algorithm (SCBA). There is no counter-reasoning in the preselection stage and no postselection stage in AI

selecting entity is separate from the entity that does the reasoning, then counter-reasoning is fruitful and, in some cases, mandatory to prepare the relevant choices for selection process. This is why the existence of counter-reasoning on its own is an indicator that the entity that ultimately selects the choice—“I”—is an independent entity from the entity that prepares the choices using reasoning and counter reasoning—the mind.

2.1.2 Selection stage

The preselection process in our mind sets the stage for the selection process in “I.” The selection stage is the heart of the decision-making process and is the main difference in the decision-making process between our mind and AI. The selection process in AI is based on SCBA and therefore, the best choice is selected based on the algorithm that AI follows. In some instances, it would be possible to add a random selection algorithm and to allow AI to choose one of a set of choices randomly in the selection stage.

However, the way we make decisions is fundamentally different. After our mind presents the choices, “I” does not follow any set algorithm for selection among the choices, and it selects a choice that would be more appealing to “I” based on ABCS. But how could “I” separate itself from any algorithm-based decision process? This would only be

possible if “I” is fully aware of the information presented in the matrix of intelligence by the mind. Only when “I” is aware of all the details of the intelligence provided for decision making can it depart from any algorithm that is set by the mind. For example, you gave your friend Jack directions to drive to the airport that include a particular highway. That is like an algorithm that you set for Jack to follow. While watching TV, another friend, Adam, sees multiple car crashes on the same highway and calls Jack and alerts him to the situation. Jack’s awareness allows him to depart—if he desires—from the algorithm that you set for him. Therefore, awareness allowed him to make a decision that is different from the original one. Similarly, awareness provides “I” with the luxury of having discretion in decision making—ABCS.

Not every decision we make is processed in “I” or involves our awareness. There are decisions that are made in our mind solely through an algorithm-based decision-making process similar to SCBA in AI. For example, the majority of our actions in sports or driving are made on autopilot, i.e., they are automatic. These decisions are made mainly in our unconscious or at a subconscious level, based on the information provided by our mind and without the need for decision-making at the level of “I.” However, it would still remain in our complete sense of autonomy to exercise our ABCS and modify or interrupt an autopilot action.

Awareness not only makes ABCS possible, it allows us to appreciate the sense of free will. We are aware of the process of decision making and all of its stages. That awareness lets us rightfully feel our discretion. In 1983 Benjamin Libet and colleagues showed in an experiment that the intention to move a limb lagged behind the brain's electrical activity indicating readiness for the movement of that limb (Libet et al. 1983). His experiment created momentum for a deterministic point of view claiming that our intention to move a muscle is nothing but a mere illusion since it appears when the move is already in action. However, since the intention to make a decision is actually the awareness of a decision made by "I" rather than the decision itself, it could be easily explained why our awareness of a decision lags behind the ABCS itself in "I."

2.1.3 Postselection stage

This is another stage in decision making that is unique to our mind and is completely absent in AI. After ABCS in "I" and completion of the selection process in the decision-making process, the selected choice is presented back to the mind for execution. Since there is a complete independence between "I" and mind, the decisions that are made in "I" are not necessarily those that will be executed in the mind. The choices that are made by "I" will go through another administrative process called appropriation that is similar to but separate from the reasoning and counter-reasoning process in the preselection stage of the decision making.

The appropriation process is an analytical mental process, like reasoning, but its main function is to check the practicality, feasibility, and executive aspects of a decision prior to its implementation. We may also call this process a reality check by our mind. If during the appropriation, the mind encounters any paradox or logical obstacle to the decision made by "I," the decision is either abandoned completely or is sent back to "I" for reconsideration. The postselection stage itself will be the subject of our awareness by "I." During reconsideration, the choices are usually further limited or modified by our mind based on real conditions on the ground. After resubmission of the reconsidered choices, "I" has the discretion to select among the modified choices or alternatively, to abandon the decision altogether.

There is no doubt that the intelligence or its processing that feeds the appropriation process is different from the reasoning process. This is why a choice that has already been selected by reasoning could be rejected by appropriation. For example, the appropriation process may find the selected choice by "I" to be incompatible with our beliefs, moral values, emotions, or capabilities. If the results of both analytical processes were the same, there would be no need for the appropriation process. It seems that while informational intelligence provides major input to the reasoning

process, emotional intelligence offers the main input to the appropriation process. For example, even when you know you made the best decision, your self-esteem may not let you move a muscle.

Obviously, there would be no reason to have a postselection stage in SCBA in AI. The same is true with autopilot decisions that are made in our mind. After a selection of the best choice by AI, the choice is executed without any ifs, ands, or buts.

2.2 Awareness process

Awareness is a translation of objective information gathered in our mind into a subjective experience. During this translation sensation, knowledge, memory and emotions transform into perception (qualia), knowing, remembering and feeling, respectively. In general, it is widely accepted that awareness is a privileged property of living minds. However, since awareness is a subjective process, it would be difficult if not impossible to rule out the existence or lack of awareness in AI, simply by an objective examination. The different stages of awareness process and their distinction in our mind and AI are depicted in Fig. 3. As I elaborate in the next part of the manuscript, our discretion plays a crucial role in our awareness process. The importance of our discretion in the selection of the information that will be processed for our awareness is fundamental to our individuality and world perspective.

2.2.1 Preselection stage

In the preselection stage of the awareness process, our mind does all the legwork to prepare the information that is presented to "I" for selection. During this preparation, our mind gathers information from various sources, including sensory input, motor information, visceral sensation, knowledge, memories, reasoning, emotional information and other forms of data and transforms them into intelligence that will be presented to "I" for the selection stage of awareness.

2.2.2 Selection stage

After preparation of the intelligence in mind, the intelligence will be presented to "I" for selection. The selection stage is key step in the awareness process. In the selection stage, only a selected information is chosen to reach the transformation stage of the awareness process based on a discretionary selection. We may call this process discretionary attention. If not for attention, our awareness would be flooded with an uninterrupted stream of information from our mind. There is no question that attention is mandatory for awareness. Even in AI, there should be a selection of particular information for processing. This may be called selection of information

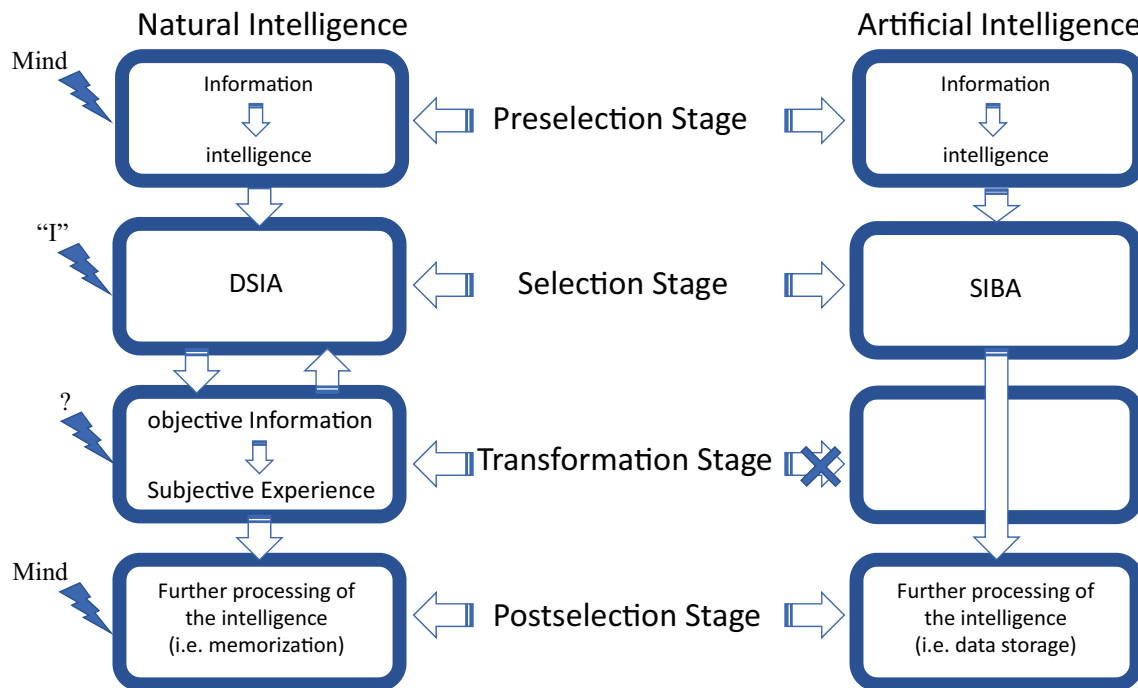


Fig. 3. Similarities and differences of the awareness process in natural intelligence (NI) versus artificial intelligence (AI). In NI we have discretionary selection of information for attention (DSIA) before the

transformation stage of awareness while in AI we have selection of information based on algorithm (SIBA) for processioning

based on an algorithm (SIBA). But why does our attention need a discretionary power—DSIA—for its selection? Could this selection be based on an algorithm? If it was not for discretionary attention, there would be no differences between individuals in paying attention to different aspects of events. Only a discretionary selection in “I” would make each of us different from other individuals. Not all the information that reaches our awareness goes through our DSIA. We do have pre-set algorithms for bypassing the selection stage for the awareness process similar to SIBA in AI. For example, a loud sound, a bright flash of light or any sudden change in sensory or motor input can reach our awareness without discretion as part of a startle reaction.

The other important fact of our awareness is the singularity of the process. That means that we can only be aware of one task at the time. We all know people are capable of multitasking. The art of switching the DSIA for awareness back and forth to different subjects in “I” could allow some people to accomplish multiple tasks using our singular awareness process. One of the distinctions between the human mind and a computer could reside in this process. Unlike our mind, AI can truly process multiple streams of information during SIBA at the same time and can literally multitask.

2.2.3 Transformation stage

After the selection of specific intelligence during the selection stage in “I,” the information goes through the transformation stage where objective information transforms into subjective experience. There is no scientific evidence to explain how or where this process takes place. Numerous theories and explanations have so far failed to provide a reasonable explanation for this process, called the hard problem of consciousness. Even though trilogy explains different stages of the awareness process, it does not cover the hard problem of consciousness. Whether or not this transformation happens in “I” is not relevant to trilogy as long as the end result of awareness process will be accessible to “I.”

2.2.4 Postselection stage

After transformation stage, the information could be selected for further processing based on our discretion. The information may now be compared with prior memories and knowledge for patterns, used for our thinking and reasoning or saved in short- or long-term memories. Similarly, in AI, processing the selected information could trigger further processing such as comparison with prior information or saving the information, all based on a preset algorithm.

At a superficial level, “I” in the trilogy may resemble a homunculus. Homunculus, or a small humanoid in the brain,

is a small human in our brain that makes decisions and was a popular notion from the sixteenth century that is generally unaccepted because of the placement of its theoretical standing on infinite regression (i.e., there must be another smaller humanoid in the homunculus and so on). In contrast, the trilogy is only the combining of the two mental processes of ABCS and DSIA, which are the core components of decision-making and awareness processes, respectively. By applying awareness to the selection stage of the decision-making process, ABCS is possible and by applying the decision-making process to the selection stage of the awareness process, DSIA is possible. These two mental processes are interdependent on each other and their dependance on each other make the other process possible. At the same time, the independence of these two mental processes from all other mental processes is crucial for their function (as it will be elaborated later). This is why this joint process that transforms our AI-like mind into a natural intelligence that we call human could be considered distinct from all other mental functions and deserves a special name—"I."

3 Trilogy vs. other consciousness theories

Trilogy does not consider consciousness interchangeable with awareness, unlike most consciousness theories. In trilogy, consciousness is an amalgam of awareness and decision-making processes. Neither awareness nor the decision-making process could possibly exist independent of each other. Below, I have presented a few relevant theories in the field of consciousness and free will to point out the similarities of and distinctions between trilogy with these theories.

Global workspace (GW) theory (Baars 1988) explains how integrated information that reaches GW will reach conscious awareness. If the information does not get the opportunity to reach GW and stays outside the spotlight, it remains in our unconscious mind. Similar to trilogy, GW is the venue for awareness, and only the information presented to GW—similar to "I" in trilogy—will reach awareness. However, this theory does not explain how and why specific information reaches GW while other information does not. Trilogy clearly spells out that the selection process based on DSIA allows the spotlight of attention to shine on a particular piece of information and bring it to our awareness.

Integrated information theory (IIT; Tonini et al. 2016) looks at consciousness from a totally different perspective. Based on this theory, integration of information results in consciousness for the entity that processes the integration. The degree of complexity of the information determines the level of consciousness of that entity. AI is as conscious as a living mind as long as the information that is being processed is as complex as the living counterpart. In IIT, similar to the other theories of consciousness, awareness

and consciousness are used interchangeably. As I explained above, in trilogy, awareness is only a part of our consciousness. Further, in IIT, there is no specific description of what integration of information actually means. If integration of information equals bundling various pieces of information, then this would stand in sharp contrast to the processing of information in the different stages of the awareness in trilogy. In trilogy, the selection of information for the purpose of awareness includes disintegration of information in the form of unbundling the information. The selection of a particular bit of information from a bundle of information based on DSIA is the key for awareness in trilogy. In trilogy, awareness is an "all or none" phenomenon, and the degree of complexity of the information does not determine the degree of awareness or consciousness. It should be mentioned that, based on the complexity of the mind and its processing capabilities, different living creatures would have awareness of simple versus complex information. Trilogy also excludes the possibility of consciousness for AI, as I argued above.

What is intrinsically missing in IIT similar to GW theory is how one piece of information go through integration while another piece of information does not. Additionally, where do we draw the line for complexity of integration that determines when a particular information reaches the level of consciousness? We know that information is being processed in different parts of our brain constantly. However, what makes one integration process more complex than another to reach our consciousness. This is where trilogy steps in and explains how one piece of information is selected over the other one for awareness based on DSIA. This selection is not due to the fact that one piece of information is more complex or further integrated than the other, but simply because of our discretion.

Recurrent processing theory (Lamme 2006) claims that a recurrent activity in a particular portion of the cerebral sensory cortex in the form of integration of information results in consciousness. This theory bridges GW theory and IIT. However, it has the same issues as its main parent theories.

Higher-order theories of consciousness (Rosenthal 2002) explain that consciousness is only possible with higher-order representation of perception when the information is presented as a first-order state. In other words, when there is a sense of self in the information, then it will reach our consciousness. Similar to IIT, integration of information is necessary for consciousness in these theories. However, only a particular integration with the sense of self can help gain information access to our consciousness. These theories also suffer from the same weakness as IIT in showing how a particular set of information can be integrated in first order while another set cannot. In trilogy, awareness is not exclusive to first-order information, and any piece of information can reach awareness. "I" does not have any sense of self on its own and is selfless.

Attention schema theory (AST; Graziano 2015) describes how consciousness is nothing but attention in our mind as well as a schema of attention. The schema of attention gives our brain the ability to have a nonphysical essence of awareness (subjective experience). In AST, there is no clear line between attention and awareness or consciousness. AST may be an interesting attempt to address how awareness can turn objective experience into subjective experience, but it falls short of explaining how we turn our attention to something and or away from something that would reach our awareness. The mere presence of this selection process makes the living mind distinct from AI, which can still be programmed to pay attention to select information based on a preprogrammed algorithm—SIBA.

4 Trilogy vs. free will perspectives

There are three major perspectives in regard to free will: libertarianism, determinism, and compatibilism. At first glance, trilogy could be considered to be aligned with the classical libertarian perspective of free will (Ginet 1966, 2008) and, in particular, agent-causal libertarianism (Chisholm 1966). In the majority of agent-causal libertarian theories, there is a metaphysical agent that makes it possible to allow detachment from the chain of cause and effect in the decision-making process. This agent is usually interpreted as a metaphysical mind or soul. However, ABCS presented in trilogy is a mental process that in combination with another mental process—DSIA—can operate within the boundaries of a physical world of causality by maintaining the independence of the choice processing within the mind from the choice selection within “I.”

But in reality, trilogy is aligned the most with the classical compatibilism perspective of free will (O’Conner 1993). Similar to the classical compatibilism point of view, determinism becomes an essential part of the decision-making process by setting up deterministic constraints in our choices (Dennett 1984; Frankfurt 1969; Kane 1999). In trilogy, we not only have a deterministic limitation in the preselection stage of the decision-making process, when choices are being formed during the reasoning and counter reasoning, but we also have deterministic constraints in the appropriation process in the postselection stage of the decision-making process. In trilogy, free will is the power to have ABCS in the selection stage of decision-making in “I” between the edge of the deterministic constraints of preselection and postselection stages of the decision-making process in the mind.

5 Trilogy vs. monism and dualism

Because the mind of an individual is not something we can observe or feel, it is considered by many as an entity separate from the body. This concept formed the root of the dualism of body and mind. The convolutions of a dualistic approach to mind and body take many forms but appear worldwide as one of the major topics that thinkers and philosophers deal with. In many religious practices, the duality is in the shape of body and soul.

Is trilogy another philosophical ideology or school of thought similar to monism and dualism by adding a third wheel to dualism? One may ask why do not we call it trialism? Trilogy is an idea but not a system of idea or philosophical school of thought. It neither is an attempt to fill the gap of dualism by bridging the gap between mind and matter. Trilogy explains how the mind works and whereas the teaching of dualism and monism could not be combined, the teaching from trilogy could be applied to either school of thought of dualism or monism. Based on trilogy, awareness is not equivalent to consciousness but only one component of it. Consciousness is the combination of awareness and decision-making process. This is because, awareness can only exist in light of DSIA, and ABCS can only be possible by the existence of awareness. Trilogy explains how a true “free will” could be possible with a deterministic backdrop without resorting to a metaphysical mind or misrepresenting a “random will” offered by indeterministic properties of quantum mechanics as a “free will”. By adding an independent entity to our mental function, trilogy allows NI to regain its discretion in the decision-making process by complete disengagement of the choice-selecting entity—“I”—from the choice-processing entity—the mind. In trilogy, awareness and independence of “I” from the mind allows detachment from all preset algorithms in the mind. Even though “I” is the master of mind, the independence of the decision-making process within “I” from the decision-executing process within mind, explains why some decisions that are made by “I” never sees the light of day. This can explain why a decision that is made in AI—unlike NI—will be executed without ifs, ands or buts. Moreover, trilogy is not simply adding a new cognitive faculty to dualism or monism but is offering “the one and only” cognitive faculty—“I”—to the mind regardless of whether one holds a monistic or dualistic point of view.

6 The similarities between our mind and AI

Now that we have elaborated the functions of the mind and AI, it seems that there is a close similarity between our mind and AI, aside from “I” (Fig. 4).

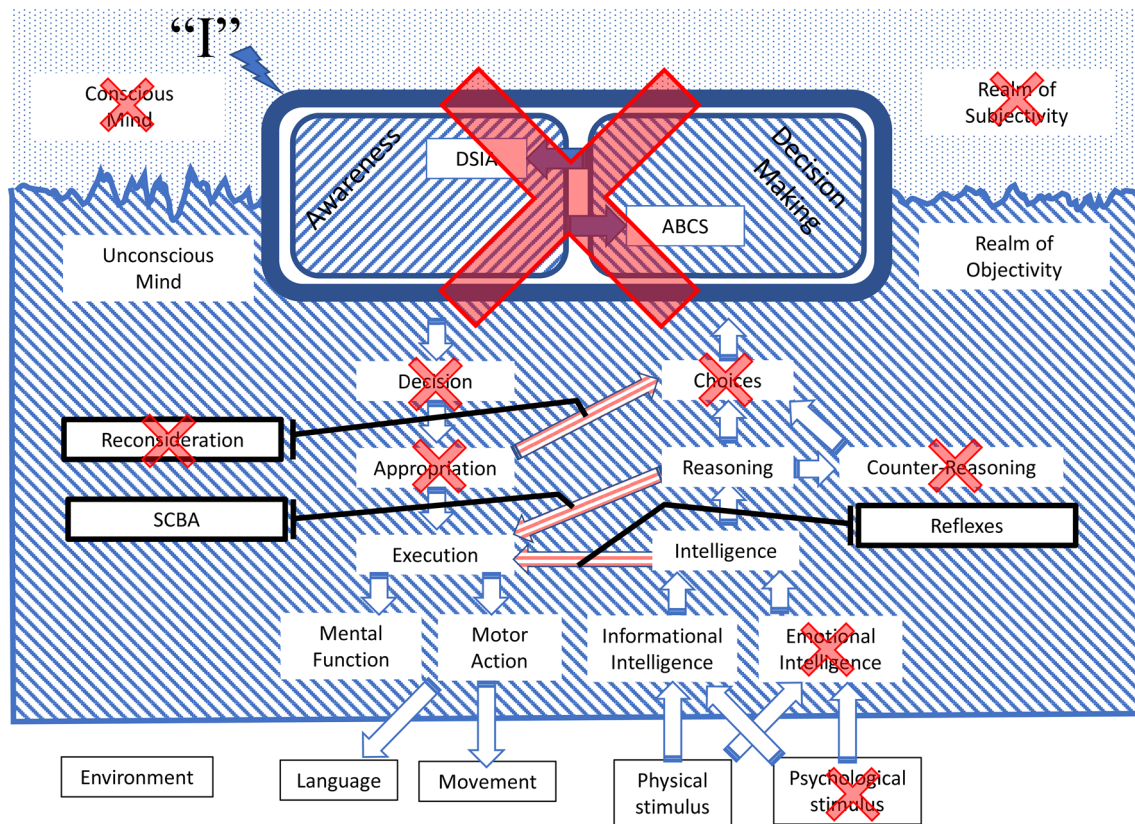


Fig. 4. The overall process of decision-making in our mind and where it differs from AI. Those processes that are marked with crosses do not currently exist in AI

Similar to our mind, AI can learn and accrue information comparable to knowledge, experience and memories. AI can also process and analyze information, identical to our reasoning. Finally, AI can make a decision based on the information in accordance with the provided algorithm. There is still a significant difference in the function of AI and our mind but some of these gaps could be easily bridged using proper programming. For example, AI does not make counter-reasoning while analyzing the information. It seems that the reasoning process that allows determination of the best choice is sufficient for making the decision that will be executed. However, this property could be easily programmed so AI could come up with multiple options in the preselection stage. However, even in that case, the selection process would be based on a provided algorithm—SCBA.

Since current AIs do not have emotions, they do not incorporate emotional intelligence in the process of reasoning. However, it would be possible to create emotional input based on common scenes similar to human encounters and incorporate that into the processed information. For example, your computer may be programmed to reply to your requests after midnight thusly: “I have been working so hard today and I am tired now. Let’s complete your Google search early tomorrow morning!”.

Similar to AI, our mind can bypass ABCS while doing autopilot actions or bypass DSIA in startle reaction. In these situations, our mind works fairly similar to AI and proceeds with SCBA or SIBA, respectively. Even in those conditions, our mind has the discretion to change the parameters of the algorithm on the fly. For example, we can change the startle sensitivity up or down or change the parameters of an autopilot action at any time based on our discretion.

7 The independence of “I” and mind in trilogy

The separation of “I” and mind in trilogy is intended solely to protect the sovereignty of “I” as a functional unit independent from the mind and all mental functions. This sovereignty allows a proper function of ABCS in the decision-making process and DSIA in the awareness process. I will explain this further in an example. If one’s collective intelligence is a deterministic engine (reasoning based on intelligence) for making a decision, and the movement of the wheels are the finalized decisions to be executed, the ABCS process in “I” is a stick-shift transmission gearbox. In this example, awareness is the clutch that

allows a detachment of the driving power of the engine and selection of a choice of a gear—ABCS—in the transmission gear box before transmitting the driving power into the wheels' motion (decisions). The independence of "I" from the engine by the clutch action allows a discretionary selection of the choice of a gear in the transmission box. Because "I" is completely independent yet aware of all informational and emotional intelligence in the mind, "I" has the luxury of making an awareness-based discretionary selection of choices that is not like the algorithmic-based, automatic decision that are made in the mind—SCBA.

The other aspects of the decision-making process could also be considered supporting evidence for the independence of "I" from the mind. For example, the mere presence of counter-reasoning in the preselection stage of the decision-making process in the mind indicates that the choices prepared in the mind are meant to be presented to an entity separate from the mind in the selection stage of the decision-making process. Otherwise, the best choice would have been enough for decision-making.

The postselection stage of the decision-making process is another indicator of this separation. Just because a decision is made by "I," that does not mean that the decision will definitely be executed by the mind. For example, you may decide to go and talk to another person but a lack of self-esteem could nip that decision in the bud before its fruition. If the decision-making entity and the decision-executing one were the same, there would have been no need for an appropriation process after a decision is made.

Further, the functional independence of "I" from the mind can preserve the integrity of the selection stage of awareness—DSIA. "I" prevents awareness process from being flooded with all the information available and allows only the selected information to reach our awareness.

"I" and mind are separated in trilogy for a functional purpose rather than because of their nature or location. There is no question that "I" functions closely with the mind at the interface of subjective and objective experiences, or at the interface of the conscious and unconscious mind. However, its functional independence from other modules of mind is a key point in trilogy. Whether "I" is physically present in a special part of the brain or is just another functional module in mind with sovereignty over other modules is merely a semantic distinction that does not change its functional importance in creating consciousness.

8 The distinction of "I" and self in trilogy

Another field of importance in the similarities and differences between our mind and AI is the topic of self-identity and self-awareness. Turing predicted that a machine may not become the subject of its own thought. Even though AI

may hold a full autobiography of itself in its memory, since it has no awareness, there is no reason to believe that AI may have self-awareness. Similar to AI, our mind is devoid of any form of self-recognition without awareness. Self-awareness is a collaborative function of "I" and mind. The information of "self" is preserved in the memory and sensory input of the mind but will only result in self-awareness when this information is presented to "I" and it reaches the awareness. If not for "I," we would not have self-awareness.

Interestingly, at first glance, "I" is the presentation of self. This is because whenever there has been a description of "I" in the history of philosophy, "I" typically has been presented as a physical or a metaphysical sense of the true self. One may ask why trilogy choose to name the independent entity, in charge of our awareness and decision making "I". Why trilogy should use this well-established entity and redefine it in a different manner from what we used to know. The reason trilogy picked the name "I" for its major player is to bring it to the attention that how the sense of self in us is not equal to the information of self stored in the mind. Based on trilogy, a sense of self-awareness is a cooperative effort of the mind which presents information about the self to our awareness in "I". In other words, all the information about the self (that could also be replicated by a computer) would never result in the sense of self without "I." At the same time, because "I" on its own does not have any self-identity without information provided by the mind, "I" is selfless.

9 Limitations of trilogy and future directives

Trilogy is based on introspection and logical deliberation. Therefore, it may not go through rigorous scientific objective examinations. Many theories of mind are solely based on introspection and subjective interpretation, but that fact does not diminish their value or negate their valuable contribution to our understanding of how the mind works. Further, trilogy provides some testable predictions based on the presentation of various stages of decision-making and awareness. Understanding these stages of decision-making could open a new horizon in behavioral science.

Trilogy is not an answer to the hard problem of consciousness. The trilogy stops where the hard problem of consciousness starts. According to the theory of trilogy, we make choices based on ABCS and we become aware of what we want to be aware of based on DSIA. This selection process is just before the transformation stage of awareness process (i.e., the hard problem of consciousness). Therefore, the trilogy would stand regardless of the physical or metaphysical perspective of the hard problem of consciousness. In other words, trilogy leaves the answer to the hard problem of consciousness to other theories and explanations.

10 Summary

In summary, there is little doubt that artificial intelligence can gather and process information in a way that's fairly similar or even superior to our thinking mind. What sets our mind apart from an AI is an entity independent from the mind called "I" that redeems our discretionary decision-making power independent from a deterministic principle of causality without a need for a metaphysical entity. "I" is composed of an intertwined process that combines our awareness and decision-making processes in two coupled mental processes of ABCS and DISA that make us conscious beings. This is in contrast to counterpart processes of SCBA and SIBA that are currently functioning in AI, respectively. This property of "I" positions it in the interface of the conscious and unconscious mind along the border of the realms of subjectivity and objectivity. In contrast to the common notion that "I" is equal to true self, "I" in the trilogy only provides awareness to the information on self provided by the mind, and this collaborative process results in self-awareness. In that regard, "I" is indeed selfless. Since "I" makes all the decisions for the mind, it represents a selfless master of the mind and is the key element that is currently lacking in AI.

Acknowledgement I want to thank the insightful comments and feedbacks offered by Dr. Shahriar Shadkhoo and Professor Patrick grim during the preparation of this manuscript.

Funding There was no funding source for this research and manuscript.

References

- Albuquerque J, Deshauer D, Grof P (2003) Descartes' passions of the soul: seeds of psychiatry? *J Affect Disord* 76:285–291. [https://doi.org/10.1016/s0165-0327\(02\)00104-0](https://doi.org/10.1016/s0165-0327(02)00104-0)
- Antonakou EI, Triarhou LC (2017) Soul, butterfly, mythological nymph: psyche in philosophy and neuroscience. *Arq Neuropsychiatr* 75:176–179. <https://doi.org/10.1590/0004-282X20170012>
- Baars BJ (1988) *A cognitive theory of consciousness*. Cambridge University Press, New York
- Bennett MR (2017) Development of the concept of mind. *Aust N Z J Psychiatry* 41:943–956. <https://doi.org/10.1080/00048670701689477>
- Chalmers DJ (1995) Facing up to the problem of consciousness. *J Consci Stud* 2(3):200–219
- Chisholm R (1966) Freedom and action. In: Keith L (ed) *Freedom and Determinism*. Random House, pp 11–40.
- Colman W (2017) Soul in the world: symbolic culture as the medium for psyche. *J Anal Psychol* 62:32–49. <https://doi.org/10.1111/1468-5922.12278>
- Dehaene S, Kerszberg M, Changeux J-P (1998) A neuronal model of a global workspace in effortful cognitive tasks. *Proc Natl Acad Sci USA* 95:14529–14534. <https://doi.org/10.1073/pnas.95.24.14529>
- Dennett D (1984) *Elbow room: the varieties of free will worth wanting*. MIT Press, Cambridge (ISBN: 0-262-04077-8)
- Dennett D, Kinsbourne M (1992) Time and the observer, the where and when of consciousness in the brain. *Behav Brain Sci* 15:183–247. <https://doi.org/10.1017/S0140525X00068229>
- Dreyfus H (1972) *What computers can't do*. MIT Press, New York (ISBN 978-0-06-011082-6)
- Facco E, Lucangeli D, Tressoldi P (2017) On the science of consciousness: epistemological reflections and clinical implications. *Explore (NY)* 13:163–180. <https://doi.org/10.1016/j.explore.2017.02.007>
- Farhadi, A. (in press). *Trilogy theory of mind: Introducing a selfless master of mind*.
- Frankfurt HG (1969) Alternate possibilities and moral responsibility. *J Philos* 66(23):829–839
- Ginet C (1966) Might we have no choice? In: Keith L (ed) *Freedom and determinism*. Random House, pp 87–104
- Ginet C (2008) In defense of a non-causal account of reasons explanations. *J Ethics* 12(3/4): 229–237. Retrieved August 23, 2020, from <http://www.jstor.org/stable/40345380>
- Graziano MS, Webb TW (2015) The attention schema theory: a mechanistic account of subjective awareness. *Front Psychol* 6:500. <https://doi.org/10.3389/fpsyg.2015.00500>
- Kane R (1999) Responsibility, luck, and chance. *J Philos* 96:217–240
- Katz L (2009) Robo-scientist makes gene discovery-on its own. *Crave-CNET News.cnet.com*. Retrieved 2010–07–29.
- Lamme VA (2006) Towards a true neural stance on consciousness. *Trends Cogn Sci* 10(11):494–501. <https://doi.org/10.1016/j.tics.2006.09.001>
- Libet B, Gleason CA, Wright EW, Pearl DK (1983) Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain* 106(3):623–642. <https://doi.org/10.1093/brain/106.3.623>
- Lucas JR (1961) Minds, machines and gödel. *Philosophy* 36 (137): 112–27. <http://www.jstor.org/stable/3749270>.
- Markoff J (2009) Scientists worry machines may outsmart man. *NY Times*, July 26, 2009.
- O'Connor T (1993) Indeterminism and free agency: three recent views. *Philos Phenomenol Res* 53(3):499–526
- Rosenthal DM (2002) How many kinds of consciousness? *Conscious Cogn* 11(4):653–665. [https://doi.org/10.1016/s1053-8100\(02\)00017-x](https://doi.org/10.1016/s1053-8100(02)00017-x)
- Russell SJ, Norvig P (2003) *Artificial intelligence: a modern approach* (2nd ed). Prentice Hall, Upper Saddle River. ISBN 0–13–790395–2
- Saraswathi TS (2005) Hindu worldview in the development of selfways: the "Atman" as the real self. *New Dir Child Adolesc Dev* 109:43–50. <https://doi.org/10.1002/cd.136>
- Searle J (1980) Minds, brains and programs. *Behav Brain Sci* 3:417–457
- Shute C (1961) The dilemma of determinism after seventy-five years. *Mind* 70: 331–350. <http://www.jstor.org/stable/2251621>
- Tononi G, Boly M, Massimini M, Koch C (2016) Integrated information theory: from consciousness to its physical substrate. *Nat Rev Neurosci* 17(7):450–461. <https://doi.org/10.1038/nrn.2016.44>
- Turing AM (1950) Computing machinery and intelligence. *Mind* 236:433–460. <https://doi.org/10.1093/mind/LIX.236.433>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.