

Article

Not peer-reviewed version

Building the Blocks of Being: The Attributes and Qualities That Are Independently Required and Jointly Sufficient for Consciousness

[Izak Tait](#)^{*}, Joshua Bensemann, Trung Nguyen

Posted Date: 1 March 2023

doi: 10.20944/preprints202303.0010.v1

Keywords: consciousness; cognition; theory of consciousness; meta-consciousness; meta-cognition



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Building the Blocks of Being: The Attributes and Qualities that Are Independently Required and Jointly Sufficient for Consciousness

Izak Tait *, Joshua Bensemann and Trung Nguyen

The NAOInstitute, the University of Auckland, 38 Princes St, Auckland CBD, Auckland, New Zealand, 1010; josh.bensemann@auckland.ac.nz (J.B.); trung.nguyen@auckland.ac.nz (T.N.)

* Correspondence: i.tait@auckland.ac.nz

Abstract: For consciousness to exist, an entity must have prerequisite characteristics and attributes to give rise to it. We explore these “building blocks” of consciousness in detail in this paper, which range from perceptive to computational to higher-order characteristics of an entity’s cognitive architecture. We show how each cognitive attribute is strictly necessary for the emergence of consciousness, and how the building blocks are collectively sufficient for any entity to be classified as being conscious. The list of building blocks is not limited to human or organic consciousness and may be used to classify artificial and organisational conscious entities. We further explore a list of attributes that seem intuitively necessary for consciousness, but on further investigation, are neither required nor sufficient. The building blocks do not represent a theory of consciousness but rather a meta-theory on the emergence and classification of consciousness.

Keywords: consciousness; cognition; theory of consciousness; meta-consciousness; meta-cognition

1. Introduction

A difficulty at the intersection of philosophical and psychological investigation into the nature of consciousness is understanding when one can properly attribute consciousness to an entity. This is in contrast to research on how consciousness is expressed and has evolved [1–3], how it correlates to neuronal and behavioural activities [4–6], or how it functions [7,8]. Our goal in this paper is to complement existing literature on consciousness by identifying the attributes and characteristics required for the attribution of consciousness.

The building blocks of consciousness are an entity’s neural, behavioural and mental characteristics that allow it to generate consciousness [3,9–11]. A *building block* means something necessary for consciousness to exist within an entity (regardless of what that entity may be). Below is an exhaustive list of building blocks that we consider sufficient together for an entity to generate consciousness.

We will not favour any specific theory of consciousness (TOC), as the building blocks below focus on the entity’s attributes rather than the nature of its consciousness. Thus, the paper will not strictly be about consciousness but what is a prerequisite to obtaining it. While we will endeavour to be as neutral and unbiased as possible regarding the theories of consciousness, please note that not all theories of consciousness will be compatible with our proposed building blocks, particularly the more heterodox theories such as panpsychism [12,13], and obviously theories that deny the existence of consciousness [14].

To expand our neutral stance, we aim to be as unbiased as possible about what sorts of entities can have consciousness. While the predominant consensus in the literature points only to humans and certain animals having consciousness [15], it is not beyond the realms of possibility for other entities to have been overlooked or prematurely dismissed. With the accelerated speed of research into machine consciousness, it may not be long before we will have to seriously consider whether an AI model is conscious or not [16]. For that, these building blocks may act as the requisite milestones that an AI model must reach before we can ascribe consciousness to it.

Most speculative yet, should we ever be (un)fortunate enough to discover complex extraterrestrial life (or if they discover us), we will need a framework to guide us to discern if they exhibit consciousness or not. Thus, we believe these building blocks will also offer some guidance to future astrobiologists.

Before we truly begin, it behoves us to provide a working definition of consciousness that will form the basis for the building blocks below. To loosely paraphrase [7], we would define consciousness as a global, holistic, and subjective mental state encompassing various conditions that have phenomenal character and cognitively functional aspects. Together, both local and global states provide an entity with the ability to be aware of, and respond behaviorally to, its internal and external environment.

2. The Building Blocks of Phenomenal Consciousness:

In this section, we look at each building block independently and discuss why each is required for consciousness yet is not sufficient by itself. Each subsection has two parts: an argument supporting the inclusion of the building block in the table and a short review providing evidence from academic literature.

The building blocks are summarised in Table 1 through lay-descriptive and narrative examples for ease of understanding.

The building blocks are not ordered in, or intended to be, a hierarchy. Much like the famous Danish toy blocks, these building blocks may be ordered in any arrangement to form the foundation of consciousness. While we argue that all the building blocks below are necessary and sufficient for consciousness, we make no claim to the form they take to do so. In addition, we do not imply that consciousness will spontaneously appear should an entity have all the building blocks below, and we will not hazard a guess as to the means by which it will do so. Those are the domains of theories of consciousness and are beyond the scope of this paper.

Note that whenever the term “environment” is used to describe the consciousness’s environment, this does not refer to the environment around the body in which the consciousness resides. Rather, it refers to everything that is not the consciousness but which interacts with it. This means that the brain is also treated as the consciousness’s environment. We don’t intend this to be an argument for dualism or against physicalism but rather a convention for ease of communication.

Table 1. Attributes and characteristics that are required for the development of consciousness.

Attribute	Description and Example
Embodiment	Information and its processing exist somewhere. <i>I am within my body.</i>
Perception (intero and extero)	Information enters from the environment and is made available for processing. <i>I see an apple.</i>
Directed attention	Information is given specific attention. <i>I pay attention to the apple.</i>
Recurrent Computing and Processing	Information is processed in and between several regions of the cognitive architecture. <i>My brain processes the image of an apple.</i>

Ability to create Inferences	Missing and incomplete information about the environment is generated. <i>I infer that it is a Granny Smith apple.</i>
Working Memory	Information is transiently maintained in an active state while it is processed. <i>I have been watching the apple for several seconds.</i>
Semantic understanding	Computing and sensing processes are understood. <i>I am aware that I am looking at an apple.</i>
Data output (internally or externally)	Information is generated and made available for further perception and action. <i>I think about apples.</i>
Introspection, meta-cognition, meta-representation	The cognitive architecture interrogates and investigates itself. <i>I think about me looking at the apple.</i>

2.1. Perception

1. To have phenomenological consciousness is to have a subjective experience of the environment.
2. This requires information exchange from the environment to the consciousness.
3. Ergo, consciousness requires a method of perceiving information in the environment.

Consciousness, whether as a phenomenological experience or as an act of attending to a matter, is defined by its interaction with the environment and receiving information from it, even if the nature of this relationship may be disagreed on [17]. To be conscious is to be conscious of something; phenomenal consciousness is about experiencing something.

Perception can be classified into three separate modes and three overlapping stages. The three modes are:

- Exteroception: perception of the environment outside the consciousness's body or housing. In humans and animals, this is best exemplified by the five classic senses; in machines, this can be microphones and cameras; and for organisations, this is the methods of communication allowed in it.
- Interoception: perception of the environment outside of the consciousness, yet within its body or housing. For animals, this includes hunger/satiety and proprioception; for machines, this may be sensors ensuring their housing and power supply is at optimal conditions; and for organisations, this may include custodial and HR oversight.
- Introspection: perception and examination of the consciousness's own existence. Section 2.8 will cover this in greater detail; however, in animals, introspection can be displayed via thoughts and mental images; in machines, via software that scans other software; and in organisations, through auditing processes.

From these three modes, a consciousness has a perceptive experience of the whole of its existence. It can perceive itself, its embodiment, and the environment around it. All three modes are not required for an entity to be considered, but at least one is. A person who, through an unfortunate accident, loses access to his exteroception wouldn't suddenly lose his consciousness.

But how are these modes perceived? According to Audi, this is done in three parts or stages, which together cause the phenomenological experience [18]. The first part is the simple perception of a scene (if visual). "I see an apple," would suffice. The second part is to perceive something within

that scene “to be” a certain way or having a certain attribute. When seeing the hypothetical apple, I can perceive it “to be” spherical even if I can only see one angle of it. When seeing its shadow, I can perceive that there is a leaf hidden from view, as the shadow seems “to be” a leaf-shaped shadow. The final part is perceiving “that” the object is what you have perceived it “to be”. I have seen an apple and perceived it to be spherical and having a leaf; therefore, my experience of it is as a spherical, leaf-having apple. When all three stages of perception are taken together, they give rise to the cognitive processes that lead to phenomenal experiences.

The final stage of perception, of perceiving “that” something is what it seems “to be”, is most crucial to consciousness. By integrating all the characteristics of the perceived object, the final stage creates a single binding and unified perceptive experience [3,10]. I do not merely perceive the sphericalness and redness of the apple, but I perceive that it is an apple. It is this final converged and singular experience which gives rise to qualia and consciousness.

But you should not take these perceptive experiences to be exclusively for the external environment. Interoception is as much, if not more, important to realising consciousness. The holistic sensory experience (conscious or unconscious) of internal signals from the body or housing plays a vital part in establishing a subjective first-person perspective for the consciousness [19]. Sensing its own body/housing allows the consciousness to differentiate itself from the rest of existence. Even if the consciousness should merely exist as Dennett's brain in a vat, it will be able to introspectively perceive and attend to its own thoughts and perception of them, causing a subjective first-person experience to emerge [4]. The mind itself becomes the environment surrounding the consciousness.

2.2. Embodiment

1. *Consciousness is affected by its environment.*
2. *A physical environment requires a physical element to this causal relationship.*
3. *This relationship, and the experiences produced by it, occurs at specific times in specific locations.*
4. *This means that the connection between the consciousness and cognitive architecture is localisable in both time and space.*
5. *A spatial and temporal physical embodiment is required for a definable perspective point.*
6. *A unique definable perspective point is required for a subjective, first-person perspective.*
7. *Ergo, an embodiment is required for consciousness.*

While a definition of consciousness rarely includes terminology of embodiment, it is a presupposition upon which many of consciousness attributes and characteristics rest [20,21]. Whether consciousness is classified as an epiphenomenon or as a complex set of phenomenological qualia, it is uncontroversial to say that consciousness is affected by its environment (except to adherents of panpsychism). There is a cause-and-effect relationship from the environment to the consciousness (and perhaps vice versa, depending on the TOC [22]) that is both spatial and temporal in nature.

For this to be true, as the logic above shows, there must be some form of embodiment for the consciousness to inhabit. From a philosophical perspective, this remains true regardless of where you position yourself on the spectrum between dualism and physicalism. For physicalists and materialists, the embodiment of the consciousness is treated as the default. Wherever the brain may go, so must the consciousness, as they are one and the same. For dualists and idealists, to whom consciousness is metaphysical, and the relationship between it and the brain is tenuous, that relationship must still terminate in a physical location (the brain), and the consciousness can only gain information via this relationship.

If you broaden the concept of consciousness to machines and AI, the notion of embodiments becomes less abstract. The lines of code from which a potential artificial consciousness must be stored on some processing and memory unit. Whether housed within one machine or spread across the virtual cloud, it is still somewhere specific in space and time.

The concept of embodiment becomes far more nebulous when looking at the potential for organisations to be called conscious entities. An organisation is irreducible; it is not its people or structures, but instead is the holistic whole that is greater than the sum of its parts. As such, the

embodiment of an organisational consciousness is a social and memetic construct. Its embodiment is located where people believe it is. Most often, this can be a physical location, particularly for traditional organisations, but for multi-national institutions and internet-based communities, the embodiment is located at the confluence of their constituents' social network, rather than in the bricks and mortar of any one building.

The number of input/output modalities of a given consciousness does not go against the principle of requiring an embodiment but does add multiple facets to this embodiment. A machine consciousness is most easily imagined as having multimodal inputs/outputs, as a cloud-based consciousness can interact and be affected by the environment through any number of terminals and devices. An organisational consciousness can likewise interact with its environment through any of the conscious entities that form part of it. Neither of these situations speak against the embodiment concept but show how embodiment can be fractured and spread across multiple locations.

Daniel Dennet's riveting essay "Where am I?" shows how multiple modalities can cause these unusual facets to the consciousness's embodiment. In the essay, his brain is removed and placed in a life-sustaining vat while remote-controlling his brain with electronic assistance [23]. Throughout the essay, he experiences his environment within the vat, in his original body, and in a second body. These modalities change, yet they (and he) are always somewhere at some time, subjectively moving from "body" to "body", yet always having a physical point in spacetime with which to access the environment.

Another important argument for why embodiment is a requirement for consciousness is that of the first-person perspective [24]. To have a subjective experience from the first-person point of view, one requires an embodiment from which to view the world. You need to be able to reference your place in spacetime in relation to the environment and be able to categorise yourself as distinct from the environment [4]. Whether through the proprioceptive, somatic and other sensory input of a biological entity, the virtual tagging of a machine, or the social identity and labelling of an organisation, an embodiment allows the entity to make itself distinct, unique, and subjective.

2.3. Directed attention

1. Entities are not consciously experiencing their entire environment simultaneously.
2. Conscious experience of the environment is limited to specific scenes, objects, or views.
3. This requires a means to discriminate between scenes/objects/views in the environment.
4. This discrimination method is achieved by directing attention to specific scenes, objects, or views.
5. Ergo, consciousness requires directed attention.

Attention is how an entity limits the information it gives priority resources to by selectively focusing on a specific selection of incoming information [10,25]. This directed attention may be triggered by a bottom-up process whereby an external event attracts an entity's attention (e.g., a flicker of movement, a distinctly recognisable sound) or through a top-down approach whereby the entity chooses to focus on a specific event or region externally or internally [3,26].

It is important to note that attention is not solely directed at the external environment, or at the interoceptive space within a consciousness's embodiment. Introspection, as discussed in Section 2.9, is defined partly by the entity's ability to direct its attention at its own cognitive processes. Attention may thus be directed at all the modes outlined in Section 2.1, from an ice cream truck's jingle heard externally, to the internal feelings of hunger, to the thoughts and memories of childhood and its connections to the present interoceptive and exteroceptive stimuli.

In this short hypothetical example above of phenomenally experiencing an ice cream truck, the observer's attention moves from extero- to interoception to meta-cognitive. This, perhaps egregiously controlled example, shows that to be conscious of each element in the phenomenological experience, you must first pay attention to that piece of information. Attention is the mechanism by which certain information passes from unconsciousness to consciousness, and other pieces of information do not. A person is clearly not attentive to every photon entering the eye or every sine wave entering the ear when one walks down Main Street. When observing the ice cream truck and reminiscing about his

childhood, the observer could very well not have been aware of a cat sitting quietly in his peripheral vision, and why should he have? His attention was focused elsewhere, after all.

While we argue that attention is a prerequisite for consciousness, the reverse is most definitely not true. One can attend to information, and direct one's attention, without it ever entering conscious processing [27]. The archetypal exemplar of this is sleepwalking. Those who suffer severely from somnambulism do not merely walk while asleep, but can do household chores, get dressed, eat, and even leave the house to drive while being entirely unconscious. They attend to the world around them, and direct their attention to what they are doing, but when they awake, they have little to no memory of anything they have done. It is attention without consciousness. There is even more than one famous case of a sleepwalker committing murder, which, one presumes, would require a great deal of attention.

More benign and ordinary instances of unconscious attention are when one's environment becomes routine enough to stop focusing on it. Dust on a bookshelf, for example, or a box of old clothes that one ought to have donated months ago. Such ordinary things may affect one's behaviour without ever consciously being aware of them [28,29], such as absentmindedly wiping off dust when walking past a shelf, or needing to adjust where one walks to not walk into that old box of clothes. Habits are formed as conscious awareness becomes unconscious attention.

One can thus frame consciousness as not attending to information received, but rather being aware of the cognitive representation of information that the architecture has already attended to [30]. In this sense, it is the interplay between working memory (holding the information that has been attended to unconsciously), the meta-representation thereof, and the recurrent processing done in the cognitive architecture that gives attention its conscious awareness [9,31]. This awareness of a mental representation of a subset of external information provides the consciousness with a view of the world, which it can phenomenally experience.

Of curious note, however, is that current AI models require directed attention because their computational resources are limited by hardware concerns (much as our cognitive architecture is). Theoretically, given enough hardware, an AI model (potentially a conscious one in the future) may not need to specifically direct its attention, as it will have enough computational resources to attend to its entire environment (internally, externally and introspectively). In such a scenario, we would argue that it is not a lack of attention, but a totality of attention that the AI model has. It had outgrown the need to focus and narrow its attention, but can encompass its entire environment in one, holistic directed attention.

2.4. Recurrent Computing and Processing

1. Conscious experience involves complex activity in several regions of the cognitive architecture.
2. Information moves recurrently between these specialised regions of the cognitive architecture involved in the conscious experience.
3. If no recurrent complex processing is detected, there are no reports of conscious experience.
4. Ergo, recurrent computational processing is required for consciousness.

We consider recurrence as another of the building blocks of consciousness. One reason for this is that there has to be an element of recurrent activity for consciousness to exist within the brain. Without any recurring activity, it would be difficult to do complex computing as the cognitive processes required to do the processing would need to complete within a single pass of the brain. However, by introducing some form of recurrence, the cognitive processes can persist as long as required to complete that processing [32]. This point is not unrelated to the importance of working memory that we have made elsewhere.

Empirical support for our position comes from the fact that many examples of recurrent activity in the brain have been documented in the literature, several of which are related to conscious processing [33]. While this does not prove that recurrence is strictly necessary for cognition, it does suggest that recurrence is essential to how the human brain produces consciousness.

Generally speaking, recurrence can take two forms for maintaining information in the brain. One is persistence, where localised activity is maintained to keep information available, such as the

process suspected to be involved in working memory [34]. The other is where information is shared between areas of the brain. This means that the information can be routed to multiple areas of the brain, each contributing to the overall cognitive process while keeping the information in existence.

The information-sharing version of recursion forms the backbone of one prominent theory of consciousness, the global workspace theory [35] and its completely neuroscientific version, the global neuronal workspace [36]. In these theories, the information becomes recurrent only if it passes an attention threshold and becomes identified as of importance to other areas in the brain. According to the neuronal version of this theory, anytime information crosses this threshold, a temporary connection forms between the relevant processing centres of the brain to allow this information sharing. Thus, important information becomes recurrent when sent back and forth between multiple processing areas.

In theory, a Deep Neural Network (DNN) could mimic the functioning of the GWT workspace by reverberating information being learned across multiple tasks in that network to all tasks being learned. Doing so might improve the performance of individual tasks. Research with multi-task DNNs has suggested that removing the independence between tasks can lead to improvements in the learning of those tasks [37,38]. An example of this could start by pooling initial task output into an area accessible by the separate sub-modules of the DNN to eliminate multi-task independence. Based on this shared information pool, it would then train the individual tasks to moderate their initial outputs. The information entering the shared space will not be ad hoc but rather will be optimised by punishing the network for sharing irrelevant information.

Having discussed how our recurrent processing requirement is compatible with global/neuronal workspace theories, it is of interest to briefly mention how recurrence relates to the other competing theory Integrated Information Theory (IIT) [39]. In IIT, consciousness is not an all-or-nothing concept; rather, it identifies a metric (PHI) that measures how integrated the information within a system becomes. This means that a human's level of consciousness can change over their lifetime with less integration occurring while their brain is developing (i.e. lower PHI scores) compared to when they are fully mature. Recurrence is not explicitly mentioned as being part of IIT, so the theory is silent on whether it is necessary for consciousness. Therefore, our inclusion of recurrence as a building block does not directly oppose IIT. With that said, IIT will allow PHI scores to be calculated for any source of information, not just biological brains. So, IIT can be applied to situations where recurrence does not exist.

2.5. Ability to create Inferences

1. A conscious entity has incomplete information about its environment due to the limitations of its sensory inputs.
2. To create a completed picture of the environment, the consciousness's cognitive architecture must build a representation of the environment with which to interact.
3. This representation is generated from inferences drawn from various parts of the cognitive architecture.
4. The information generated from inferences is not limited to perceptual data but may include cognitive information such as feelings.
5. This generated information is then available to be outputted as experiences.
6. Ergo, the ability to create inferences is required for phenomenal consciousness.

A consciousness never has a complete view of its environment, be it the inner world of its embodiment, the external world beyond it, or even the introspective view of its own cognitive architecture. This is because the sensors it has of its environment limit the information it receives at any one time. Our memories are not perfect, our vision has blindspots and areas of inattention, our hearing is rather limited, and our interoceptive sensors are mostly unconscious.

Yet, from a subjective point of view, we do indeed have a 'completed' view of our environment, if not entirely 'complete'. This incongruity is solved via the use of inferences by the cognitive architecture. Through the use of inferences, the cognitive architecture generates the "missing" information to provide the consciousness with a unified view of its environment. This generative act

of building inferences is not merely limited to perception but can be modelled throughout the cognitive processes, even to a meta-cognitive level.

Note that inferences in this section do not imply causal inference in the way that a cognitive architecture is estimating and reasoning a cause behind an event. Rather, One can think of the basal inferences involved in perceptual events referred to here as analogous to the interpolation of statistical data.

One can look at inferences from a hierarchical perspective. At the “lower” levels, closest to the sensory inputs, would limit their inferences on what is being perceived to create a unified perceptive view [40] and build the meta-representations of the environment. Higher levels further removed from the sensorium would then continue the meta-representational process by generating inferences on the holistic view of all inputs. Throughout these sensory inferences, the cognitive architecture would build its “best guess” of the environment, continually updated through sensory information to reduce any errors in its predictive guesswork [41].

This predictive inferential model works equally well in memories. Biological memories are notoriously unreliable, partly due to the inferences worked into them. As memories are recalled (consciously or unconsciously), additional information is generated to fill in any missing areas to construct a narratively enjoyable memory. This can often occur due to cues in the present environment that colours the perception of the memory and the inferences generated [42]. This generated information can then become part of the memory when it is next recalled and then reiterated upon with further inferences until all that is left of the subjective past is imagination.

Inferences are not purely perceptual in nature. Subjectivity, feelings and emotional states also have inferential components. In biological creatures, emotion relies heavily on interoceptive inferences of the physiological changes of the embodiment [43]. Yet, from a conceptual point of view, an entity’s “feeling” about a given matter can be seen as the inferences made of the differences and juxtapositions between the internal and external environment. Perceptions of the external world are correlated with the physiological state of the internal environment by generating predictions of how the two should interact and correlate. As the physiological internal state changes and as the perceptions of the external world change, these predictions no longer hold true and “prediction errors” crop up. Inferences of an emotive nature are generated to reconcile these prediction errors that give rise to subjective feelings and, in turn, a phenomenological experience.

The subjective first-person experience is also generated by this reconciliation of exteroceptive and interoceptive environments [43,44]. As an entity develops, it builds an understanding of what and where it is in relation to the environment by generating inferences. This is as much a social as it is perceptual. By perceiving that others exist, the entity infers that it is different from them. By perceiving the external world and the limitations the entity has in its interactions with it, it can infer that it is not the same as the environment. As no aspect of the internal or external environment is static, the predictive model in the entity’s cognitive architecture is also constantly updated to generate new inferences about itself and its environment [45,46].

An example of this is expected changes over time. As we see ourselves and those around us ageing (particularly if we have not seen someone in a significant time), we view new iterations of ourselves as others, rather than the same exact view from our memories. We infer who they are, and who we are ourselves when we look in the mirror, through the predictions that our cognitive architectures have built. Despite minor changes in appearance, there is a continuity of experience through the use of active inferences [44].

2.6. Working Memory

1. A specialised unit or process is required to maintain transient information as it is being processed in various regions of the cognitive architecture.
2. Working memory is responsible for holding and maintaining said information.
3. Ergo, working memory is required for consciousness.

Working memory (WM) is considered a cognitive system that stores, maintains, and processes online (short-term) information relevant to the very immediate/current task [47,48]. The maintained

information in WM is about what is currently being thought about and experienced to be consciously considered. Therefore, it is arguable that some processing of conscious experience occurs here. Specifically, there are several general views on the relationship between WM and consciousness as they are considered equal or closely working together [31].

The first popular view is that WM is closely related and implicitly considered equal consciousness. In other words, the content held in WM is consciousness, as shown in Baddeley's multicomponent model [47,49]. This WM model is hierarchical, with a central executive and several "child" systems, where the central executive controls the slave systems. The child systems are either modality-specific information or the episodic buffer of polymodal episodes [49]. While the "child systems" provide consciously experienced representations, the central executive may relate to consciousness by providing conscious access to items held in WM. Many theorists assumed this view in the recent past. This idea, however, is hard to stand with the evidence of unconsciousness content in WM.

Another view is to consider consciousness as an integrated component of WM which enables WM to contain unconscious content. This idea was demonstrated in the embedded-process WM model [50], activation-based models of WM [51,52] and the opinions of having a distinction between conscious WM content and unconscious WM content (which may become conscious) [53]. In this view, selecting WM content as the focus of attention could be the primary mechanism mediating the relationship between WM and consciousness as it activates original WM content to enable conscious access to the content [31].

The second view is favoured by [54] argument that WM representations and conscious representations serve different functions, have different effects on behaviours [55–57] and have distinct representations [58,59]. Therefore, [54] proposed a third view that introspected contents are the "conscious copy" of WM representations. This model suggests that WM representations are intrinsically unconscious, and consciousness is not the same as the WM trace, even if the WM content is activated. Nevertheless, conscious experience is rooted in WM in this theory and, therefore, requires WM to exist. [54] also mentioned another opinion that WM is a subset of consciousness, opposite to the second view. This view again supports the requirement of WM for consciousness, although it requires further empirical evidence.

The relationship between consciousness and WM is an ongoing research topic with divergence. Although the relationship can be either hierarchical or parallel interaction, WM is integral in the workflow of consciousness in all models and general views of this research. WM's equivalence have been applied and integrated into artificial computational systems, from computer hardware to complex models. In computer hardware, RAM is considered to be equivalent to WM, fulfilling many of the same requirements and processes, with introspected memory extracted from peripherals and local/cloud data. In Franklin's implementation of the global workspace theory (GWT) [60,61], i.e., Intelligent Distribution Agent [62], WM is considered equivalent to the whole global workspace, which supersedes consciousness. What this may mean for potential organisational conscious entities is that the people who make up the organisation plays the role of WM, with their own memories and consciousnesses working to maintain the transient information in the organisations communications networks.

2.7. Semantic understanding

1. Subjective perceptual awareness differentiates conscious from unconscious experience.
2. Awareness of a specific scene, object or view in the environment requires understanding that the process is occurring.
3. To be aware that an experience is subjective, the entity must understand that it exists in some capacity.
4. Ergo, semantic understanding is required for consciousness.

There is a fundamental difference between processing information and being aware of it. Haikonen (2020) describes the distinction between the two types of processors when discussing Weak AI versus Strong AI. In Weak AI, information can be input into a system, processed, and then output

without the processor having any experience of those processes. There are many examples of processing without experience in neuroscience (see Dehaene & Naccache, (2001)). For AI, the information input into the system is transformed into a number, making it indistinguishable from any other process occurring in the AI. In other words, the Weak AI is limited by the Symbol Grounding Problem.

To achieve a Strong AI, Haikonen (2020) argues that the Symbol Grounding Problem must be solved so that the AI has some form of experience of the outside world that is different from converting inputs into numbers. While previously arguing that this problem might be impossible to solve with digital computers [63], Haikonen does explain what successfully solving the problem looks like, using Humans as an example. Essentially, the problem is solved if the processor has some kind of experience about the information it is processing, such as qualia experienced in humans.

Qualia are one of those constructs that have many conflicting definitions. While there are many different points of view on what qualia are [64], we will briefly outline one from [65]) to help explain what it is the above paragraph describes. Ramachandran and Hirstein provide three criteria that must be satisfied for things to qualify as qualia. The first is that the experience is irrevocable: The agent experiencing the quale cannot use cognitive effort to change what they are experiencing. For example, blue always appears as blue; we cannot choose to see it as red.

The second criterion is that the agent must have some choice about how they react to the experience. An example of this is the difference between how a coma patient and an awake individual react to light being shined into their eyes. The coma patient may only constrict their pupils via a reflex, so they are not consciously experiencing the light. In contrast, the normal individual experiences this because they can close their eyes, turn their head, complain about the light, or do any other action.

The third criterion is that the experience needs to enter working memory. As discussed before, working memory is one of our building blocks. If a stimulus does not enter working memory, then it is detected and processed without the agent's awareness. Without awareness, the experiencer does not actually experience anything and so cannot be experiencing qualia. Consequently, if the experience does not enter working memory, it is impossible to determine whether that experience is irrevocable. Additionally, if the experiencer is unaware of the stimulus, they cannot factor in that experience into their decision-making and so fail to meet the second criterion. As the first and second criteria are impossible without working memory, the third criterion appears to be the most important for determining whether something is experiencing qualia.

The above definition is just one description of what is required for something to be an experience. We do not commit to it but have included it as a useful description of what semantic understanding can be. The example also helps highlight the importance of working memory as a building block. However, we should point out that the semantic understanding building block is a specialised task and is distinct from working memory. One should also compare the view we have included here to others, such as Reggia et al. (2016), who further discuss experience from a self-modelling perspective. This is a much larger topic that we will not expand upon here.

2.8. Data output

1. Subjective experience is a crucial component of phenomenal consciousness.
2. This experience includes perceptive and/or phenomenal elements such as thoughts, emotions, feelings, words and actions.
3. These elements do not come from sources external to the cognitive architecture.
4. These perceptive elements must thus be created by the cognitive architecture to be subjectively then experienced.
5. Ergo, data output is required for consciousness.

Crucial to the definition of phenomenal consciousness is the experience of an event or scene from a subjective first-person perspective. This experience of "what it is like to be" at that time, at that point, within that embodiment, localised entirely from that perspective, is not an empty event devoid of characteristics. It is the characteristics and attributes of the experiences which allows them to be

felt by the consciousness. This means then that some form of information generation is required to create these experiential characteristics.

Three obvious and intuitive types of experiential information may first come to mind, but they are not truly required for consciousness. Yet, for those who can generate them, these play an important role. These are verbal/aural and visual mental imagery and emotional information. Section 3 below will detail the reasons why these are not required for consciousness due to disorders such as anauralia, aphantasia and alexithymia. External behaviours are also not required informational content for consciousness, as individuals with severe paralysis or pseudocomas can still show evidence of consciousness and mental responsiveness without being able to show it externally.

This leaves one prominent piece of information that can be generated regardless of known disorders and is vital to the definition of a phenomenal experience: a feeling. As discussed in Section 2.5 above and Section 3 below, feelings are distinct from emotions in that they do not require a physiological response and encompass a broader spectrum of content than simply emotions. Emotions may also have unconscious aspects to them [66], while feelings need to be consciously felt. For a consciousness to 'feel' anything, the cognitive architecture must generate that 'feeling' in some form or fashion.

Even in dualistic and idealistic theories of consciousness, there is no feeling at one moment, and then at the next moment, a feeling is spontaneously there. As feelings are inherently subjective and will differ from entity to entity when experiencing the same event, we can confidently say that these feelings are not provided to the consciousness from the environment external to the consciousness's embodiment. The envatment notion put forth by Dennet in 1978 also shows that interoception is useful yet insufficient for a phenomenal feeling. Therefore, even if an anti-materialist stance is taken, the cognitive architecture is the only conduit between the physical realm and the consciousness where a feeling may be generated or felt.

Data generation is also vital to other building blocks, as both introspection (Section 2.9) and inferences (Section 2.5) have elements that require the creation of information. As such, data output plays both a causal and affected role in establishing consciousness.

2.9. Introspection, meta-cognition, meta-representation

1. Higher-order processing of perceptive computational elements is required to form a subjective experience.
2. Higher-order processing involves creating mental representations of the environment.
3. These representations are a recursive process whereby the cognitive architecture creates mental representations of mental representations, known as meta-representation.
4. Higher-order meta-representation involves the consciousness experiencing its own experiencing, or thinking about its thought processes, known as introspection or meta-cognition.
5. Ergo, introspection, meta-cognition, and meta-representation are required for consciousness

Introspection, meta-cognition and meta-representation all have the common element of one section of cognitive architecture creating representations of another, "lower" section.

Introspection may be the most obvious as it is the subjective examination, inspection and perception of the cognitive architecture by the consciousness [67]. We know that we are conscious because we can think about being conscious. Or as Descartes famously said, "Je pense, donc je suis": I think, therefore I am [68].

Meta-cognition is broader still; it is any cognitive process that is about another cognitive process rather than the embodiment's external environment. This can include thinking about the details of a memory rather than simply recalling that memory, or judging the effort and difficulty of a task based on knowledge of one's own skills. One can think about meta-cognition as a hierarchical and recursive process, where the cognitive architecture investigates and interrogates itself around and around from higher sections to lower sections and back again [69].

Meta-representation, to follow the naming convention, is thus about the cognitive architecture creating representations of other representations. While intuitively related to meta-cognition, meta-

representation is distinct in that it does not require directed thought, and the representations may be of the cognitive architecture itself or the external environment [70].

With these three aspects together, one can build a narrative of phenomenal experience as it is perceived. Presuming, as in the table at the start of this document, that one sees an apple. The first-order section of your cognitive architecture would process this image, and you would, indeed, mentally see the apple. Yet, only when attention (Section 2.3) is directed towards the apple, and understanding of this directed attention (Section 2.7) is had, would you start to be conscious of the apple. Recurrent processing (Section 2.4) is thus already active as multiple sections of the cognitive architecture work together to do this.

Merely seeing an apple means nothing if your cognitive architecture does nothing with it. As the information is passed through different sections of the cognitive architecture, secondary sections of the architecture become involved. These sections are not strictly tied to perception and may fill several purposes, and as they are a step removed from the external environment, they can be called higher-order sections. To make sense of what is being seen, these sections would create mental representations of the visual representations, guided through what is in working memory (Section 2.6) and inferences about the environment (Section 2.5).

At this point, the architecture is fully conscious of the apple. It can create further representations of the apple by drawing from memories (if it had seen anything like an apple before) or on its logical processes to determine what it may be. These are cognitive processes about representations of the apple, not the apple itself, which means they are meta-cognitive. As one ruminates on the idea of the apple, thoughts and feelings are generated (2.8), which may lead to introspection and further consideration.

This narrative example may have been lengthy, but should the first step have been the memory of an apple rather than perceiving a physical one, this process would be truncated in half and begin with meta-cognition and introspection rather than exteroception.

The argument against first-order processing being sufficient for phenomenal consciousness is two-fold. Firstly, first-order perception (of any kind mentioned in Section 2.1) is predominantly unconscious. Memories are often recalled without intention, and any moving object may catch the eye to wake one from a daydream. The unconscious mind is extremely powerful and processes far more than one would think, but consciousness requires an additional step; it requires one to “pay attention” [71]. One section of the cognitive architecture needs to direct the section closest to the “event”.

The second argument is that there is a distinctly different subjective experience of external phenomena based on one’s own architecture that is separate from the sensation itself [4]. For example, if you understand English, reading this sentence will provide a different subjective feeling and experience than if you could not understand English.

The simple act of meta-cognition and introspection can itself generate the cascade that leads to phenomenal experience. If one were to think, “Am I forgetting something?” then one would generate the content which would be perceived by one’s own consciousness, creating meta-representations and meta-cognitive processes required to explore and evaluate that thought, accompanied by feelings and experiences and the train of thought rolls along. This shows how introspection and meta-cognition are not only unconscious processes, but consciously directed ones that can modify other conscious or unconscious processes [67].

One should not take the above reasoning to imply that introspection, meta-cognition, and meta-representations at the human level and quality are required for consciousness. We know that young toddlers and infants do not have full meta-representations of the external environment [70], yet they do have phenomenological consciousness. Similarly, gaining introspective reports from animals can often be difficult, if not impossible. One can argue for degrees and scales of introspection and meta-cognitive abilities, and that as long as the minimum level required is reached, the entity would be classified as conscious.

3. What is not required for consciousness:

Some aspects of what makes us who we are are intuitively linked to our consciousness. Table 2 has a selection of these that, when first given thought, we would like to say are necessities for consciousness as they seem so inextricably linked to it. However, there are neurological and psychological conditions wherein each item in Table 2 is either missing or severely reduced. As individuals with any of these conditions still display all other signs of consciousness, we can confidently state that they are not absolute necessities for the generation of consciousness, as the building blocks are in the previous section.

Table 2. Attributes and characteristics that seem intuitive to consciousness but are not required nor sufficient for its existence.

Attribute	Disorder in which it is deficient
Emotions	Alexithymia, various Schizoid Personality Disorders, Depression, Psychosis, Emotional Detachment.
Inner Speech	Anauralia.
Inner Visualisation	Aphantasia.
Theory of Mind	Infancy, Autism, Bipolar Disorder, Schizophrenia, Attention Deficit Hyperactivity Disorder.
Long-term Memory	Amnesia (retrograde and anterograde).

The first item on the list may seem the most intuitively linked to consciousness. Crucial to having a subjective first-person experience is having the feeling of “what it is like to be”. Commonly, this experience has some emotional attachment, but not always. Despite the two terms often being used interchangeably, there is a distinct difference between feelings and emotions. At best, emotions can be called a subcategory of feelings, in that emotions require a distinct physiological state change, while a feeling can be purely mental. In this sense, the concept of “feelings” is much broader than that of “emotions”.

There are several disorders and illnesses where emotions may be subdued or missing. Emotional detachment (particularly in cases of psychopathy, but also linked to trauma) and depression may be the most common conditions where there is a reduced range and strength of emotional expression. In various schizoid personality disorders and most famously in alexithymia, the strength of emotional responses is put on a spectrum, where individuals may only feel a sense of subdued emotions, all the way through to not being able to label or identify any emotional reactions they may be having. The absence of emotion, however, does not seem to hamper their ability to have subjective experiences.

Conceptually, both machines and organisations lack the physiological requirements to produce an emotional reaction. At best, we will need to program a machine to mimic emotion, while only the individuals within an organisation have emotions. Yet, organisations can display all other signs of conscious experiences, and, theoretically, so too will machines one day.

While the name for the next condition is very recent, there is a disorder in which an individual does not have inner speech, called anauralia [72]. These individuals have no running internal monologue (or dialogue), and no thoughts are “spoken aloud” inside the mind. Most often, people with this condition are unaware of it until they discover that others can indeed speak to themselves inside their heads (apologies to those reading this and discovering your condition for the first time).

This latter point is most important for why inner speech is not required for consciousness. As many individuals do not know that they have a disorder, they continue their lives in blissful ignorance, having as rich and full subjective conscious experiences of the world as those with inner

speech. Their lack of verbal thoughts does not hamper any of the building blocks above, and they can verbally report on introspection and subjectivity equally as well as others.

On the other side of this coin is aphantasia, a condition where one's mental ability to imagine a visual object or scene is placed on a spectrum. Aphantasia is entirely independent of the lack of inner speech [73] yet follows a highly similar pattern. At the higher degrees of this spectrum, individuals with aphantasia have difficulty generating a mental image, while at the lower degrees, there is a complete absence of mental visual creation. As with those with little to no inner speech, there is no evidence that aphantasia has any significant negative impact on the phenomenal character of conscious experiences [74].

One of the most often cited requirements for, and expressions of, consciousness is the Theory of Mind, or the ability for an entity to understand that other entities have conscious thoughts and experiences different from their own. It feels so intuitive that we ought to understand that our minds are different to others to have a subjective point of view, yet we know that infants do not have a theory of mind, and most toddlers only develop these by around the age of three or four [75]. At the same time, we can see that infants and young toddlers most certainly have a subjective first-person point of view, and those with early language development can provide introspective reports on these phenomenological experiences.

In addition, those with autism, schizophrenia, ADHD, bipolar disorder, severe mental and language impairment, or those with traumatic brain injury have been shown to have a heavily reduced theory of mind [76–82]. While we must note that the degree of theory in mind in each of these medical cases is on a spectrum, taken together with the absence in infancy, it shows that it is definitely possible for humans to have phenomenological experiences while having a reduced or no theory of mind.

Last on the list is long-term memory. Working Memory is most definitely a building block of consciousness, but long-term memory is not required. As a case study, one can look to the famous example of Clive Wearing, an esteemed musicologist who gained severe retrograde and anterograde amnesia after losing a bout with herpesviral encephalitis. Wearing is continually stuck in the present, his conscious experience lasting approximately thirty seconds before “resetting” [83]. Yet, despite this, Clive still recognised his wife, even though he could not tell who she was, could still conduct and perform music (in limited amounts owing to time), and could talk about how he subjectively felt during his brief moments of “wakefulness” [84].

Despite not being able to form new episodic memories, and having lost almost all of his memories before becoming ill, Clive Wearing could still show (for thirty seconds at a time) all the semblances of consciousness that one would expect. He could attend to information, report on introspection, understand what he was doing, and even create inferences about his condition. While his condition had an undeniably devastating effect on his quality of life, it did not overly diminish his ability to have phenomenal experiences.

4. Conclusion

In this paper, we described nine cognitive features, attributes and characteristics that are each individually required for an entity to be classified as conscious. We also propose that if an entity has all nine building blocks, it would be holistically sufficient for said entity to have generated consciousness.

The purpose of this paper is not to create yet another theory of consciousness, but to serve as a guide for identifying if an entity is conscious or not. Conceptually, the building blocks apply equally to identifying consciousness within natural, artificial or organisation entities (and perhaps, one day, even extraterrestrial). Furthermore, as research advances into creating conscious, superintelligent machines, these building blocks may serve a series of milestones that ought to be reached before that AI can be classed as being conscious (while several have already been reached already).

The paper also included a short list of features that initially seemed intuitively required for consciousness but were found not to be when investigated further. The list of non-required features is as vital as the list of building blocks themselves, as it allows us to expand the range of potential

conscious entities, but also to look at consciousness from a less neurotypical anthropocentric point of view.

As mentioned in the introduction, the building blocks, by definition, are required to be an exhaustive list, and we have done our level best. However, no theory is ever written in stone. We welcome all critiques, suggestions and discussions on the building blocks and whether we ought to include more, less, or even if some need to be lumped together or split apart as the taxonomist are wont to do.

References

1. Bensemann, J.; O'Leary, P.; Chen, Y.; Miranda-Dukoski, L.; Witbrock, M. Simulations and the Evolution of Consciousness. In Proceedings of the The 2022 Conference on Artificial Life; MIT Press: Cambridge, MA, 2022.
2. Arbib, M.A. Co-Evolution of Human Consciousness and Language (revisited). *J. Integr. Neurosci.* **2014**, *13*, 187–200.
3. Feinberg, T.E.; Mallatt, J. The Nature of Primary Consciousness. A New Synthesis. *Conscious. Cogn.* **2016**, *43*, 113–127.
4. Reggia, J.A.; Katz, G.; Huang, D.-W. What Are the Computational Correlates of Consciousness? *Biologically Inspired Cognitive Architectures* **2016**, *17*, 101–113.
5. Lepauvre, A.; Melloni, L. The Search for the Neural Correlate of Consciousness: Progress and Challenges. *Philosophy and the Mind Sciences* **2021**, *2*.
6. Wu, W. The Neuroscience of Consciousness. *The Stanford Encyclopedia of Philosophy* 2018.
7. Seth, A.K.; Bayne, T. Theories of Consciousness. *Nat. Rev. Neurosci.* **2022**, *23*, 439–452.
8. Tononi, G.; Koch, C. Consciousness: Here, There and Everywhere? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2015**, *370*, doi:10.1098/rstb.2014.0167.
9. Arrabales, R.; Ledezma, A.; Sanchis, A. ConsScale: A Pragmatic Scale for Measuring the Level of Consciousness in Artificial Agents. *Journal of Consciousness Studies* **2010**, *17*, 131–164.
10. Birch, J.; Ginsburg, S.; Jablonka, E. Unlimited Associative Learning and the Origins of Consciousness: A Primer and Some Predictions. *Biol. Philos.* **2020**, *35*, 56.
11. Panksepp, J. The Periconscious Substrates of Consciousness: Affective States and the Evolutionary Origins of the Self. *Journal of Consciousness Studies* **1998**, *5*, 566–582.
12. Chalmers, D.; Others Panpsychism and Panprotopsyism. *Consciousness in the physical world: Perspectives on Russellian monism* **2015**, 246.
13. Frankish, K. Panpsychism and the Depsychologization of Consciousness. *Aristot Soc Suppl Vol* **2021**, *95*, 51–70.
14. Shenker, O. Denialism: What Do the so-Called Consciousness Deniers Deny? • *The Jerusalem Philosophical Quarterly* **2020**, *68*, 307–337.
15. Edelman, D.B.; Seth, A.K. Animal Consciousness: A Synthetic Approach. *Trends Neurosci.* **2009**, *32*, 476–484.
16. Metzinger, T. Artificial Suffering: An Argument for a Global Moratorium on Synthetic Phenomenology. *J. AI. Consci.* **2021**, *08*, 43–66.
17. Doyon, M. Perception and Normative Self-Consciousness. In *Normativity in Perception*; Doyon, M., Breyer, T., Eds.; Palgrave Macmillan UK: London, 2015; pp. 38–55 ISBN 9781137377920.
18. Audi, R. Perception and Consciousness. In *Handbook of Epistemology*; Niiniluoto, I., Sintonen, M., Woleński, J., Eds.; Springer Netherlands: Dordrecht, 2004; pp. 57–108 ISBN 9781402019869.
19. Cosmelli, D.; Thompson, E. Embodiment or Envatment? Reflections on the Bodily Basis of Consciousness. In *Enaction: Toward a New Paradigm for Cognitive Science*; Di Paolo, E.A., Stewart, J.R., Stewart, J., Gapenne, O., Eds.; MIT Press Limited, 2010; pp. 361–386 ISBN 9780262014601.
20. Nguyen, D.J.; Larson, J.B. Don't Forget About the Body: Exploring the Curricular Possibilities of Embodied Pedagogy. *Innovative Higher Education* **2015**, *40*, 331–344.
21. Glenberg, A.M. Embodiment as a Unifying Perspective for Psychology. *Wiley Interdiscip. Rev. Cogn. Sci.* **2010**, *1*, 586–596.
22. Thompson, E.; Varela, F.J. Radical Embodiment: Neural Dynamics and Consciousness. *Trends Cogn. Sci.* **2001**, *5*, 418–425.
23. Dennett, D. Where Am I? In *Brainstorms: Philosophical Essays on Mind and Psychology*; Bradford Books: Cambridge, 1978; p. 11 ISBN 9780262540377.
24. Rudrauf, D.; Williford, K. Unlimited Associative Learning and the Origins of Consciousness: The Missing Point of View. *Biol. Philos.* **2021**, *36*, 43.
25. Juliani, A.; Arulkumaran, K.; Sasai, S.; Kanai, R. On the Link between Conscious Function and General Intelligence in Humans and Machines. *Transactions on Machine Learning Research* **2022**.

26. Koch, C.; Tsuchiya, N. Attention and Consciousness: Two Distinct Brain Processes. *Trends Cogn. Sci.* **2007**, *11*, 16–22.
27. Mole, C. Attention and Consciousness. *Journal of Consciousness Studies* **2008**, *15*, 86–104.
28. Dijksterhuis, A.; Aarts, H. Goals, Attention, and (un)consciousness. *Annu. Rev. Psychol.* **2010**, *61*, 467–490.
29. Bello, P.; Bridewell, W. Attention and Consciousness in Intentional Action: Steps Toward Rich Artificial Agency. *J. AI. Consci.* **2020**, *07*, 15–24.
30. Graziano, M.S.A.; Webb, T.W. The Attention Schema Theory: A Mechanistic Account of Subjective Awareness. *Front. Psychol.* **2015**, *6*, 500.
31. Velichkovsky, B.B. Consciousness and Working Memory: Current Trends and Research Perspectives. *Conscious. Cogn.* **2017**, *55*, 35–45.
32. Funahashi, S.; Procyk, E. Editorial: Persistent Activity in the Brain - Functions and Origin. *Front. Neural Circuits* **2021**, *15*, 841451.
33. Mashour, G.A.; Roelfsema, P.; Changeux, J.-P.; Dehaene, S. Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron* **2020**, *105*, 776–798.
34. Dehaene, S.; Naccache, L. Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework. *Cognition* **2001**, *79*, 1–37.
35. Baars, B.J. *A Cognitive Theory of Consciousness*; Cambridge University Press, 1988;.
36. Dehaene, S.; Kerszberg, M.; Changeux, J.-P. A Neuronal Model of a Global Workspace in Effortful Cognitive Tasks. *Proceedings of the National Academy of Sciences* **1998**, *95*, 14529–14534.
37. Bensemann, J.; Witbrock, M. The Effects of Implementing Phenomenology in a Deep Neural Network. *Heliyon* **2021**, *7*, e07246.
38. Long, M.; Cao, Z.; Wang, J.; Yu, P.S. Learning Multiple Tasks with Multilinear Relationship Networks. In Proceedings of the Proceedings of the 31st International Conference on Neural Information Processing Systems; Curran Associates Inc.: Red Hook, NY, USA, December 4 2017; pp. 1593–1602.
39. Tononi, G. An Information Integration Theory of Consciousness. *BMC Neurosci.* **2004**, *5*, 42.
40. Friston, K. Consciousness and Hierarchical Inference. *Neuropsychanalysis* **2013**, *15*, 38–42.
41. Seth, A.K. From Unconscious Inference to the Beholder's Share: Predictive Perception and Human Experience. *Eur. Rev.* **2019**, *27*, 378–410.
42. Conway, M.A.; Loveday, C. Remembering, Imagining, False Memories & Personal Meanings. *Conscious. Cogn.* **2015**, *33*, 574–581.
43. Seth, A.K. Interoceptive Inference, Emotion, and the Embodied Self. *Trends Cogn. Sci.* **2013**, *17*, 565–573.
44. Fotopoulou, A.; Tsakiris, M. Mentalizing Homeostasis: The Social Origins of Interoceptive Inference. *Neuropsychanalysis* **2017**, *19*, 3–28.
45. Sajid, N.; Ball, P.J.; Parr, T.; Friston, K.J. Active Inference: Demystified and Compared. *Neural Comput.* **2021**, *33*, 674–712.
46. Cooke, J.E. What Is Consciousness? Integrated Information vs. Inference. *Entropy* **2021**, *23*, doi:10.3390/e23081032.
47. Baddeley, A. Working Memory. *Science* **1992**, *255*, 556–559.
48. Baddeley, A. Working Memory: Theories, Models, and Controversies. *Annu. Rev. Psychol.* **2012**, *63*, 1–29.
49. Baddeley, A. The Episodic Buffer: A New Component of Working Memory? *Trends Cogn. Sci.* **2000**, *4*, 417–423.
50. Cowan, N. An Embedded-Processes Model of Working Memory. In *Models of working memory: Mechanisms of active maintenance and executive control*, (pp 61–101); Miyake, A., Ed.; Cambridge University Press, xx: New York, NY, US, 1999; Vol. 506, pp. 62–101.
51. Oberauer, K. Access to Information in Working Memory: Exploring the Focus of Attention. *J. Exp. Psychol. Learn. Mem. Cogn.* **2002**, *28*, 411–421.
52. Engle, R.W. Working Memory Capacity as Executive Attention. *Curr. Dir. Psychol. Sci.* **2002**, *11*, 19–23.
53. Baars, B.J.; Franklin, S. How Conscious Experience and Working Memory Interact. *Trends Cogn. Sci.* **2003**, *7*, 166–172.
54. Jacobs, C.; Silvanto, J. How Is Working Memory Content Consciously Experienced? The “Conscious Copy” Model of WM Introspection. *Neurosci. Biobehav. Rev.* **2015**, *55*, 510–519.
55. Perky, C.W. An Experimental Study of Imagination. *Am. J. Psychol.* **1910**, *21*, 422–452.
56. Craver-Lemley, C.; Reeves, A. Visual Imagery Selectively Reduces Vernier Acuity. *Perception* **1987**, *16*, 599–614.
57. Pan, Y.; Lin, B.; Zhao, Y.; Soto, D. Working Memory Biasing of Visual Perception without Awareness. *Atten. Percept. Psychophys.* **2014**, *76*, 2051–2062.
58. Bona, S.; Cattaneo, Z.; Vecchi, T.; Soto, D.; Silvanto, J. Metacognition of Visual Short-Term Memory: Dissociation between Objective and Subjective Components of VSTM. *Front. Psychol.* **2013**, *4*, 62.
59. Bona, S.; Silvanto, J. Accuracy and Confidence of Visual Short-Term Memory Do Not Go Hand-in-Hand: Behavioral and Neural Dissociations. *PLoS One* **2014**, *9*, e90808.

60. Baars, B.J. *In the Theater of Consciousness: The Workspace of the Mind*; Oxford University Press, 1997; ISBN 9780195102659.
61. Baars, B.J. Global Workspace Theory of Consciousness: Toward a Cognitive Neuroscience of Human Experience. *Prog. Brain Res.* **2005**, *150*, 45–53.
62. Franklin, S. A Conscious Artifact? *Journal of Consciousness Studies* **2003**, *10*, 47–66.
63. Haikonen, P.O. Consciousness and Robot Sentience Available online: <https://www.worldscientific.com/worldscibooks/10.1142/8486> (accessed on 22 November 2022).
64. Chan, L.-C.; Latham, A.J. Four Meta-Methods for the Study of Qualia. *Erkenntnis* **2019**, *84*, 145–167.
65. Ramachandran, V.S.; Hirstein, W. Three Laws of Qualia: What Neurology Tells Us about the Biological Functions of Consciousness. *Journal of Consciousness Studies* **1997**, *4*, 429–457.
66. Kriegel, U. Towards a New Feeling Theory of Emotion. *Eur. J. Philos. Sci.* **2014**, *22*, 420–442.
67. Overgaard, M.; Mogensen, J. An Integrative View on Consciousness and Introspection. *Rev. Philos. Psychol.* **2017**, *8*, 129–141.
68. Descartes, R. *Discours de la Méthode Pour bien conduire sa raison, et chercher la vérité dans les sciences*; 1637;.
69. Peters, M.A.K. Towards Characterizing the Canonical Computations Generating Phenomenal Experience. *Neurosci. Biobehav. Rev.* **2022**, 104903.
70. Proust, J. Metacognition and Metarepresentation: Is a Self-Directed Theory of Mind a Precondition for Metacognition? *Synthese* **2007**, *159*, 271–295.
71. Lau, H.; Rosenthal, D. Empirical Support for Higher-Order Theories of Conscious Awareness. *Trends Cogn. Sci.* **2011**, *15*, 365–373.
72. Hinwar, R.P.; Lambert, A.J. Anauralia: The Silent Mind and Its Association With Aphantasia. *Front. Psychol.* **2021**, *12*, 744213.
73. Amit, E.; Hoeflin, C.; Hamzah, N.; Fedorenko, E. An Asymmetrical Relationship between Verbal and Visual Thinking: Converging Evidence from Behavior and fMRI. *Neuroimage* **2017**, *152*, 619–627.
74. Lennon, P. Aphantasia and Conscious Thought. *Oxford Studies in Philosophy of Mind* **2022**, *3*.
75. Korkmaz, B. Theory of Mind and Neurodevelopmental Disorders of Childhood. *Pediatr. Res.* **2011**, *69*, 101R–8R.
76. Capps, L.; Kehres, J.; Sigman, M. Conversational Abilities Among Children with Autism and Children with Developmental Delays. *Autism* **1998**, *2*, 325–344.
77. Happé, F.; Frith, U. The Neuropsychology of Autism. *Brain* **1996**, *119* (Pt 4), 1377–1400.
78. Senju, A. Spontaneous Theory of Mind and Its Absence in Autism Spectrum Disorders. *Neuroscientist* **2012**, *18*, 108–113.
79. Bora, E.; Bartholomeusz, C.; Pantelis, C. Meta-Analysis of Theory of Mind (ToM) Impairment in Bipolar Disorder. *Psychol. Med.* **2016**, *46*, 253–264.
80. Sprong, M.; Schothorst, P.; Vos, E.; Hox, J.; Van Engeland, H. Theory of Mind in Schizophrenia: Meta-Analysis. *Br. J. Psychiatry* **2007**, *191*, 5–13.
81. Muller, F.; Simion, A.; Reviriego, E.; Galera, C.; Mazaux, J.-M.; Barat, M.; Joseph, P.-A. Exploring Theory of Mind after Severe Traumatic Brain Injury. *Cortex* **2010**, *46*, 1088–1099.
82. Yirmiya, N.; Erel, O.; Shaked, M.; Solomonica-Levi, D. Meta-Analyses Comparing Theory of Mind Abilities of Individuals with Autism, Individuals with Mental Retardation, and Normally Developing Individuals. *Psychol. Bull.* **1998**, *124*, 283–307.
83. Rathbone, C.J.; Moulin, C.J.A.; Conway, M.A. Autobiographical Memory and Amnesia: Using Conceptual Knowledge to Ground the Self. *Neurocase* **2009**, *15*, 405–418.
84. Wilson, B.A.; Wearing, D. Prisoner of Consciousness: A State of Just Awakening Following Herpes Simplex Encephalitis. In *Broken memories: Case studies in memory impairment*, (pp; Campbell, R., Ed.; 1995; Vol. 444, pp. 14–30.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.