

# Consciousness and Subjectivity, Artificial Intelligence and Theism

## Consciousness (in its full subjective and reflexive expression) – a definition:

*“the stratified operations of a network to process information, in its own awareness and power of volition, and a subjective sense of itself, thinking and acting voluntarily or involuntarily as an independent agent, both in relation to the sensory inputs of its environment and past experiences.”*

### Explanation of terms:

- **“stratified”** because it has a number of layers of operation and associated levels of awareness. For example, to be aware of your breathing to have to make a decision to do that ... to bring what is essentially an autonomic reflex into conscious awareness. In Orthodox Christianity (and indeed differently expressed in many other religions and meditative practices) this is a preparatory feature of hesychastic prayer.
- **“network”** (unspecified) leaves open the question as to whether this could be organic or inorganic, carbon or silicon based, or running on some other form of wetware or hardware.
- **“awareness”** is the most basic aspect of consciousness, but it can still be present in differing levels of intensity and complexity. Human awareness, neurologically speaking, involves a filter and edit system with approximations and memory playing their parts in analysing a *selection* of data, the fullness of which might otherwise overwhelm us. This “fuzzy logic” can be twisted by the conundrums of optical illusion, situations where the brain cannot decide how to interpret its visual stimulus. Internally this mental awareness can, in some individuals, in a condition called synaesthesia, cross reference sensory input such that songs or numbers can also be perceived as colours, words as smells or taste sensations etc. Awareness is never uncomplicated or exact, internally within the mind or externally in relation to the environment and its sensory input. The mind can even be fooled by mental disorders or brainwashing into storing false memories of things which never happened, experiences never had.
- **“volition”** – awareness is not enough for full consciousness, there must be an ability both to think (to interact with data) and to decide, free of constraint, to do something with that input. Pain, at one level, is simply an experience that is perceived as unpleasant. This has survival value and enables the prevention of harm. If I touch a hot plate, the autonomic response of my central nervous system will quickly retract my hand. I do not have to think about that action. It just happens, (unless the nerves in my fingers are damaged). However, consciousness in relation to that pain is to learn from that experience, to remember and, therefore, to plan the avoidance of such painful actions in the future. I have to *decide* to do that as a conscious thinking subjective agent. This is reflected in St Augustine’s triad of conscious living, memory, intellect and will (volition – voluntary or involuntary). Pain, by the way is an example of what philosophers call **“qualia”**. These are subjective experiences everyone knows or feels in common, but which are exceedingly difficult to describe. Other examples include colour or musical notes.
- **“subjective” and “independent”** – these are the necessary corollaries of memory, intellect and will. To have a true sense of “self” as a subjective reality, we learn as babies to distinguish ourselves from other selves (such as ‘warm-feeding-thing-mother’) as independent subjects, which of course makes for a truly free relationship between subjects, imbued with value, positively or negatively. This sense of self in toddlers does not fully develop until the age of two!<sup>1</sup>

---

<sup>1</sup> <https://theconversation.com/how-do-children-develop-a-sense-of-self-56118>

## The 3 different levels of Consciousness

Researchers in neurology have identified 3 different levels of consciousness in ascending degrees of complexity and self-reference (*as defined by Dehaene, Lau and Kouider*).<sup>2</sup>

These classifications can be useful in beginning to answer the question:

*“Could Artificial Intelligence in the fields of computing or robotics ever achieve states of consciousness comparable to that of human beings?”*

### The three levels of consciousness:

1. **C0: Unconscious processing** – the hand retracts after accidentally touching the hot plate.
2. **C1: Transitive processing** – the hand tests the heat of the hot plate by feeling the heat above it.
3. **C2: Reflexive processing** – the mind thinks about other dangers of burning it has known (striking a match for example when young). In so doing it will retrieve from memory experiences of these past examples of C0 in learned and retained behaviours. This is the higher level of awareness known as ‘metacognition.’ Human consciousness always includes C2. The “hard problem of consciousness” is also to be found in the debate between materialists and dualists on the interface of brain and mind.

### More Resources on Consciousness as experienced by humans

[David Chalmers in interview with Lex Fridman](#) (*the first 31 minutes is speculation about reality and simulation which is not germane not to our topic and which you can skip if you wish*).<sup>3</sup>

Chalmers surveys the field fairly but by his own admission is a proponent of pan-psychism from a post-modern point of view. Pan-psychism teaches that everything has a degree of consciousness within a universal field of consciousness. Chalmers is a post-modernist in that he claims to be able to derive values and purpose, not from any universal consciousness set (pantheistic God?) but from within his own consciousness; like so many other post-Protestant Enlightenment self-made men. Nothing new there!

Wikipedia has an excellent and reliable article on [The Hard Problem of Consciousness](#) which maps out the different approaches to this issue, past and present.<sup>4</sup>

## Artificial Intelligence and Consciousness

This is the key question:

*Could Artificial Intelligence achieve Consciousness and a Sense of Self / Subjectivity comparable to that experienced by humans?*

Existing AI systems can be programmed to fulfil C0 and C1 processing functions. Collision avoidance sensor systems in cars are a good example. The risk of collision is registered by proximity sensors and these trigger an alarm to the driver and / or avoidance automatic braking. We can even program such systems to learn and adjust their responses accordingly (Bayesian algorithms). Nonetheless, whether this is wholly

---

<sup>2</sup> <https://science.sciencemag.org/content/358/6362/486.full>

<sup>3</sup> <https://youtu.be/LW59IMvxmY4>

<sup>4</sup> [https://en.wikipedia.org/wiki/Hard\\_problem\\_of\\_consciousness](https://en.wikipedia.org/wiki/Hard_problem_of_consciousness)

in the initial programming or whether the machine learns for itself, the processing is still at levels C0 and C1, even in the most complex AI systems now available to us.

No machine we have at present, even those that mimic our own neural processing in neural networks, achieves C2 processing either in the soft consciousness form of problem solving or in the hard consciousness form of self-awareness and true subjectivity. We have no idea how such machine networks and systems could acquire, emergently or by initial design alone, that form of higher-level consciousness. This is not to say, theoretically, that they could not. Faith and theism cannot rule out that possibility. Transhumanism <sup>5</sup> (generally antagonistic to all religion but arguably becoming a cult itself) argues that humans as they are now will soon become obsolete as the machines evolve far past us in the so-called “singularity” event. It is essential that the Orthodox Church develops its own theological reflection on these technologies, which many do indeed believe to present an existential threat to all humanity.

## The Theological Implications of AI with higher level consciousness (C2)

Fundamental to Orthodox Christian anthropology is the teaching that all humans are made according to the image of God (Genesis 1:26) and with the potentiality after the Fall and by the redeeming work of Christ to attain unto the full likeness of God. The Fathers have differing accounts of what constitutes this “image,” but St John of Damascus is fairly typical of them all:

Since this is so, God created man out of visible and invisible nature with his own hands according to the image and likeness, forming the body from the earth and through his own breathing upon it giving it a rational and intellectual soul, which we call the divine image. That which is “according to the image” is manifest in the intellect and free will. That which is “according to the likeness” is manifest in such likeness in virtue as is possible. ORTHODOX FAITH 2.12 <sup>6</sup>

A “rational and intellectual soul ... freewill ... virtue” certainly embrace and indeed require a C2 reflexive higher level of consciousness as described herein before. As far as we are aware, none of the other members of the animal kingdom possess this and our created machines certainly do not, as of now at least. However, we must press on to ask the (albeit) hypothetical question: *If machine AI's did acquire that higher level consciousness would that be identical in all respects to the human “rational and intellectual soul” replete with virtues, as reflecting the dependable goodness of God?* We cannot apply the Turing test to this question, a test which arguably has only limited applicability to certain features of the human mind and consciousness. However, I suppose we might be able to refine the Turing test <sup>7</sup> to cover ALL aspects of mind and consciousness, in which case, the honest answer would be: “unlikely but who knows?!”

If such human creations did become indistinguishable from their fathers and mothers would these machine children also be made in the image of God? No, but they would be made, partially at least, in our image and likeness, thus allowing for St Paul description of humans as co-workers (co-creators?) with God (1 Corinthians 3:9). Might they then “fall” - as in the speculations of Isaac Asimov? <sup>8</sup> Would they have rights - as defended by Captain Jean Luc Picard in respect of the android, Commander Data in Star Trek? Who knows? These are questions for the future; but we do need to start thinking them through now before we are overtaken by events!

---

<sup>5</sup> <https://whatistranshumanism.org>

<sup>6</sup> Louth, A., & Conti, M. (Eds.). (2001). Genesis 1–11 (p. 35). Downers Grove, IL: InterVarsity Press.

<sup>7</sup> [https://en.wikipedia.org/wiki/Turing\\_test](https://en.wikipedia.org/wiki/Turing_test)

<sup>8</sup> [https://en.wikipedia.org/wiki/Three\\_Laws\\_of\\_Robotics](https://en.wikipedia.org/wiki/Three_Laws_of_Robotics)