

# HOW MUCH CONSCIOUSNESS DOES AN IPHONE HAVE?

Gary Marcus

---

*The following originally appeared as a blog at NewYorker.com (6 June 2013). Marcus is a professor at New York University.*

What has more consciousness: a puppy or a baby? An iPhone 5 or an octopus? For a long time, the question seemed impossible to address. But recently, Giulio Tononi, a neuroscientist at the University of Wisconsin, argued that consciousness can be measured—captured in a single value that he calls  $\Phi$ , the Greek letter phi.

The intuition behind Tononi's idea, known as the Integrated Information Theory, is that we experience consciousness when we integrate different sensory inputs. According to Tononi, when you eat ice cream, you cannot separate the taste of the sugar on your tongue from the sensation of the melting liquid coating the inside your mouth. Phi is a measure of the extent to which a given system—for example, a brain circuit—is capable of fusing these distinctive bits of information. The more distinctive the information, and the more specialized and integrated a system is, the higher its phi. To Tononi, phi directly measures consciousness; the higher your phi, the more conscious you are.

Over the past few years, the theory has become increasingly influential; it has even been championed by the eminent neuroscientist Christof Koch, a Caltech professor and the chief science officer at the Allen Institute for Brain Science.

There are several reasons to take Tononi's ideas about phi seriously. Unlike many other theories of consciousness, his gives scientists and philosophers a quantitative way of grappling with the possibility that creatures like mice and cats might have some degree of awareness (though less than humans). The theory also helps explain why certain relatively complicated neural structures don't seem critical for consciousness. For example, the cerebellum, which encodes information about motor movements, contains a massive number of neurons, but doesn't appear to integrate the diverse range of internal states that the prefrontal cortex does.

An interesting consequence of the theory, at least as Tononi and Koch have articulated it, is that anything with a phi greater than zero possesses at least a shred of consciousness. By that definition, many organisms, and even some

computers, are conscious by virtue of the ways they integrate information.

At least two computer programs exist that would score a relatively high phi, yet it seems unreasonable to call either one "conscious." IBM's Watson and Google's self-taught visual system, which learned to detect cats in images simply by examining millions of stills from YouTube videos (many of which, it turns out, feature cats), would both seem to register a substantial amount of phi because they absorb vast quantities of data. But Watson lacks self-awareness, and while Google's cat detector can recognize faces and other features, it doesn't have the slightest idea of what those things mean. It would seem odd to say that it has an experience of "catness" in the way that a human does when he sees a cat. (Tweaking those programs, or simply making them more massive in scale, might ostensibly result in a higher phi, but it's not clear that it would make them more conscious.)

Meanwhile, phi is ridiculously hard to compute, making it difficult for scientists to fully evaluate the theory behind it. As Tononi says, the value "reflects how much information a system's mechanisms generate above and beyond its parts." The only way to quantify it precisely is to consider the exponentially large number of ways a neural system might be arranged, and to compare every possible whole with every conceivable configuration of its parts; the more complicated the system, the harder it is to evaluate.

The upshot is that, even though phi promises in principle to be precise, it can't actually be used in any workable sense. What is the phi value of the average human brain, with its eighty-six billion neurons? What about a cat's brain? Tononi and Koch have no idea. There is currently no practical way to calculate those numbers, because an unthinkable large number of possibilities would have to be evaluated. (It is a safe bet that the average person has a higher phi than the average cat, but without doing the insanely demanding calculations, it is hard to say exactly how much higher.)

But even if phi could be accurately assessed, a correlation with consciousness would not in itself provide proof of causation. For one thing, a phi value (or any other measure of the way information is integrated and distributed across the mind) could be merely a prerequisite for consciousness,

and not necessarily a signal of its presence. It might also be simply correlated with consciousness rather than a measure of it, as the philosopher Ned Block wrote to me in an e-mail. Block suggested that phi is actually a barometer of intelligence rather than of consciousness per se. (And, as Block further notes, consciousness and intelligence can be understood to be decoupled in principle, as in science-fiction stories with super smart machines that are not, in fact, conscious.)

To fully understand what defines consciousness, we need more than a single measure of information flow. We may need to better understand how organisms' inputs mat-

ter, how those organisms ground their experiences in the world, and how intelligence relates, causally, to consciousness itself.

We will also need to understand more about what information percolates in the brain, and where; and about what kind of computations are performed in the course of processing that information. Phi clearly gives us a new way to think about the relationship between information and consciousness, but it is probably too abstract to ever be a complete explanation of consciousness. For that, we will need to understand exactly how our brains are wired.