

From visual perception, and towards a theory of consciousness

Aravinda Korala, 25th May 2020

[Go to website](#)

This document is an attempt at formulating a theory of human consciousness that can be used to drive experiments as well as computer simulations of consciousness. I am aware that that is a high bar I am setting, but I am hoping that you the reader will help me move this forward with your feedback.

A few words about me

First a few words about me and why you should bother reading this at all. I am not a philosopher nor a researcher of AI today, but I have a PhD in Computer Vision and AI from 1985. I am a Computer Scientist with several decades of experience with Software and Hardware. Since 1989 I have been the CEO of a software company with a focus outside of AI - so I am also a businessman. I am a little out of date with the literature, but I have tried to read as much as I could online before writing this. I apologize if I missed works of importance – but I am sure someone will point that out and I will try to incorporate.

Introduction

It is clear that humans are conscious, while sticks and stones are not, and that animals are somewhere in between depending on whether they are primates or ants. The boundary of where consciousness starts is tricky for sure and depends on our definition of consciousness as much as on our ability to measure that consciousness.

A good place to start our discussion is human visual consciousness. Firstly, because it is so highly developed and is so very vivid to us. But also because we can experiment with it more easily than with say touch or taste. (But most importantly because that is what I researched in depth from 1979 to 1985).

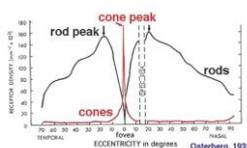
I will structure the rest of the document as statements followed by justifications or explanations below it.

1. Our visual experience begins when light from the environment enters the eye and is focussed onto the [retina](#). [Rods and cones](#) in the eye sense the light and transmit electrical impulses via the [optic nerve](#) to the brain.

An important aspect of the mechanics of the eye is that the retina is not covered uniformly with visual sensors like a modern camera is. The very central part of the retina, the [Foveola is 250 µm wide](#), has a visual angle of just 1.2° and a [cone density](#) of 147,000/mm² at its peak. However, in total the foveola has just 17,500 cones - ie colour sensors. The [foveola](#) is the central part of the fovea, and the fovea has approximately 200,000 cones and a visual angle of about 5°. The fovea and the foveola are critically important when we need the highest possible visual resolution, such as when reading a book.

This is in fact quite low resolution when compared with a modern camera. [An 18 mega pixel camera](#) has a pixel density of 1 million pixels per mm² and has 18m pixels uniformly arranged, and a visual angle that can be very wide depending on the lens that is used.

The retina *does* have more sensors than just within the fovea. This link has some [interesting statistics](#) of the human retina. The most important point I want to address here is that the visual resolution of the human eye is not at all uniform across the field of view. Check out the graphic and link below.



(Please click on the link in the image above)

The cones are the colour receptors of the eye, but notice the way the cone density spikes in the foveola, with very low density outside that central area. The rods on the other hand are colour blind, and see only in black-and-white, but are [sensitive](#) to low light as well as to movement. Notice also the way there is a blind spot where the optic nerve attaches to the retina.

The primary point I want to make here is that the sensors of the human eye are not laid-out at all like that of a modern camera.

2. Although the “camera” of the human eye has a non-obvious physical structure as shown above, our visual consciousness does not “see it” that way. We do not see a blind spot, we do not think we see in near black-and-white outside the central area, we are not conscious that our visual acuity is very poor outside the immediate 1.2° that we see with the central foveola, and we are not conscious of the fact that within that foveola area we see with just 17,500 colour-sensitive cones – a very limited resolution.

In fact, if we built a camera with the physical structure of the human retina the resulting visual image would be very poor indeed. The video signal would consist of a very small central disc of high resolution – just wide enough to contain one word at reading distance – and a low-resolution near black-and-white picture with degrading resolution outside of that central disc.

How then is our visual experience so very different from that?

3. There are two parts to the answer. The first is that the eye never stays still - it is constantly [in motion](#) “scanning” the visual field. The second is short term [visual memory](#). As the visual field gets scanned by the eye, it also gets remembered in short term memory. So the eye is more like a laser scanner that scans the scene and builds up a picture as it does so, and far less like a [CCD camera](#).

The eyes move voluntarily as well as involuntarily. The involuntary [saccadic movements](#) are thought to be critically important to our visual capability.



The picture above shows the way saccadic movements scan what the observer considers are the important parts of the image. It is not at all like an X-Y raster scan of a laser. Keep in mind though that the visual image that the eye sees is much richer than the single-pixel at a time of a laser scan. With each saccadic movement, a highly sensitive colour sensor of 17,500 sensors is aimed at the “interesting part” of the image with an additional 200,000 sensors of the fovea surrounding it. That is of course why the eye does not need to scan in an X-Y sense. It can jump to the interesting areas while seeing a significant part of the

surrounding area too. From the graphic above, it can be seen that the eyes, nose and mouth are important for humans when observing a facial image.

4. The important conclusion for me here is this: we do not perceive the retinal sensor data directly. That retinal image appears to get written into a memory “buffer” that is short term visual memory, and our visual experience is what we see from that buffer.

Computer Scientists love [buffers](#). We are buffering things all the time – there are hardware buffers and software buffers used in computing. So for example a hard disk might have a buffer that is often called a [cache](#) that might hold data waiting to be written to the disk, or data waiting to be read by the CPU. That cache may be on the hard disk itself. The operating system (e.g. Windows) may in addition cache large swathes of disk storage in its [file cache](#). Indeed, on a PC with 16GB of RAM half of that memory could be caching files temporarily so that they can be accessed quickly by the CPU when needed.

Human [visual short-term memory](#) appears to be similar to our computer buffers. As the Wikipedia article referenced above says “ the visual input consists of a series of spatially shifted snapshots of the overall scene, separated by brief gaps. Over time, a rich and detailed long-term memory representation is constructed from these brief glimpses of the input, and VSTM is thought to bridge the gaps between these glimpses and to allow the relevant portions of one glimpse to be aligned with the relevant portions of the next glimpse.”

5. This then is a first indication that human visual perception is not a “waterfall” process whereby visual sensor data pours into the eyes and flows along while being processed at each stage, until we have visual perception at the bottom of that waterfall. As a minimum, visual sensor data is being cached in short-term memory – at least for the purpose of overcoming weaknesses inherent in the visual hardware of the eye.

This triggers an obvious question. Is the visual short-term memory holding the sensor data at a rod-and-cone-level of detail – or has it already been processed to some extent so that the caching (or buffering) is at a higher level of “abstraction”? In other words, has the pixel level of data from the sensors been reduced in detail, so that it is both more useful (because “irrelevant detail” has been removed) and less bulky (in terms of data volume), and therefore easier to store?

[It appears](#) that there are in fact two types of caching. Very short term “iconic memory” that is of high capacity and operates at the level of the sensors of the eye and lasts 100ms or so, and longer term visual short term memory that lasts several seconds but holds the data at [a more abstract level](#) and is of much lower capacity and detail.

6. The primary point I want to make here is that human visual processing does not progress as a video camera system might do where a whole scene is acquired in a fraction of a second and is then processed and transmitted onwards. Instead the human eye scans the scene with saccadic eye movements and larger head movements, using what is a low-resolution camera system, stores that data in an intermediate sensory buffer to allow the scene to be built up bit by bit as an image scanner would. This image is processed to a higher level of abstraction and is stored in visual short-term memory where it can hold that image abstraction up to a few seconds, but at a lower level of detail.

The importance of this as we will soon see is that the bottom-up process can take visual perception only so far...

7. What happens next is crucial to our narrative. There appears to be a mechanism of [bottom-up and top-down](#) processing that work together, where knowledge-based top-down processing interacts with sensor-data based bottom-up processing to create an understanding of the visual content.

[David Marr](#) 1980 argues in his influential [book](#) "Vision" that there is a sophisticated process of image processing whereby low level features such as edges are extracted first, followed by higher level features such as surfaces that result in a 2.5D "primal sketch". This is then further analysed by the brain and results in a 3D model of the scene. See the table below from page 37 of Marr's book.

Table 1-1. Representational framework for deriving shape information from images.

Name	Purpose	Primitives
Image(s)	Represents intensity.	Intensity value at each point in the image
Primal sketch	Makes explicit important information about the two-dimensional image, primarily the intensity changes there and their geometrical distribution and organization.	Zero-crossings Blobs Terminations and discontinuities Edge segments Virtual lines Groups Curvilinear organization Boundaries
2½-D sketch	Makes explicit the orientation and rough depth of the visible surfaces, and contours of discontinuities in these quantities in a viewer-centered coordinate frame.	Local surface orientation (the "needles" primitives) Distance from viewer Discontinuities in depth Discontinuities in surface orientation
3-D model representation	Describes shapes and their spatial organization in an object-centered coordinate frame, using a modular hierarchical representation that includes volumetric primitives (i.e., primitives that represent the volume of space that a shape occupies) as well as surface primitives.	3-D models arranged hierarchically, each one based on a spatial configuration of a few sticks or axes, to which volumetric or surface shape primitives are attached

There are several compelling features to Marr's thesis. The idea that features are extracted from the visual image which creates an internal representation of the scene is powerful. The

idea that the brain stores those representations is powerful. The idea that existing representations then get used to analyse new scenes is very powerful. I find these concepts compelling.

However, I am not sure I agree with the intermediate feature types that Marr chose, such as edges, surfaces and geometric shapes. But that is perhaps not so important for the purpose of this paper - Marr was not necessarily saying that humans created geometric representations in their heads.

It is for instance unlikely that humans have a high-quality 3D representation of the world that can be manipulated and used as a computer might do. For instance, it is clear that we cannot easily recognize faces that are shown [upside down](#), nor are we able to recognize shapes shown from an unusual angle [unless we have already memorized that view](#).

That is not to say that our recognition is statically linked to exactly what we have seen before. So, for example, we can recognize a face through a vast range of image sizes – from a picture the size of a postage stamp to a picture that might be the size of a wall. We can also recognize a face even when it is partially occluded – but that ability falls away rapidly depending on how much of the face is still visible. On the other hand computer programs are [probably able](#) to perform better than humans with partial data of a face.

The important conclusion here is this: although human visual perception is quite incredible and is unmatched by any artificial systems we are able to build today, it is also limited in ways that computers are not. We are able to recognize and manipulate visual imagery provided it is available to us in the “normal” way - not upside down for example.

8. What is important then for this narrative is the concept that the brain has representations of objects that it has previously seen and uses that previous knowledge to recognise and analyse new scenes.

This reinforces an important statement from earlier. Visual perception is not a bottom-up process where light enters the eye and perceptions fall out the other end. It is a highly *active* recognition process whereby knowledge of previously learned scenes and objects are combined with the visual stream from the eye, along with saccadic eye movements and head movements to create a final visual consciousness.

9. Idea: here is a proposal for building a visual perception system that could potentially be as good as human vision.

- a) **Begin with low level feature extraction from the visual image. It is important to retain information about local colour gradients and intensity gradients as much as possible in these low-level representations.**
- b) **Identify areas of local uniqueness – for example those caused by rapid colour-gradient changes. (eg the same way the human visual system [locks-in](#) on the area around the eyes nose and mouth, while ignoring the cheeks and forehead).**
- c) **Combine these local areas of interest into a connected [graph](#) of local areas.**
- d) **Now observe the way this connected graph of unique areas of an object move together in unison when that object is moved or rotated – just like a baby does as she learns her visual environment. Movement of objects is powerful to visual learning in my view. It tells the visual system which parts belong to the object, and which do not.**
- e) **Use massive indexation of the local features like [Google Search](#) does. Indexation is the key to instant recall – consider how quickly Google can find a unique bit of text from 100 peta bytes (10^{17}) of data! We need to index locally unique areas in the same way and recognize objects using those indexed features.**

I published some of my early ideas about this in my [PhD thesis](#) of 1985. It focused on a small subset of objects and used a local feature detector that was perhaps somewhat eclectic, but I think demonstrated the concept.

I did not however implement my point #9d nor did I implement indexation as in #9e above.

Some notes:

- #9a – keep in mind that when extracting features, we must retain as much “information” as possible while reducing as much “data” as possible. A contradiction for sure. The “information” that matters is the information that describes the object – such as colour gradients – but independent of lighting conditions and shadows as far as possible. Manipulating and moving the object in 3D helps of course. Use of colour gradients reduces the impact of lighting, while 3D manipulation reduces the impact of shadows during the learning process.
- There is a #9f ... as Google says in the article on indexation “with the Knowledge Graph, we’re continuing to go beyond keyword matching to better understand the people, places and things that you care about.” A system that implements 9a-9e would be good – but would still struggle with describing mountains and lakes that are not conducive to being manipulated as in 9d above. Additional help is needed when learning large objects that cannot be manipulated, or that we cannot walk the system around. Even more importantly, general knowledge of the world is required before perceptual processes can compete with humans. You cannot hope to recognize the New York city skyline and understand that within the context that humans do without a significant amount of general knowledge, as Google alludes to above.

With regards the actual techniques and technology of machine vision, it is important to say that there is a huge research community and a vast amount of literature on it (see for example [Bob Fisher](#)). The field is too huge for me to review and I leave that to others. I hope my proposal above #9a - #9f adds value to that work.

10. Let's now talk about auditory perception as a stepping stone towards our journey to describe general consciousness. There is evidence that human auditory perception works in a similar manner to the human visual system.

Auditory short-term memory [appears to behave](#) like visual short-term memory. The authors Visscher et al conclude "These results imply that auditory and visual short-term memory employ similar mechanisms". And why not? Why wouldn't evolution copy what it got right with visual perception?

I do not plan to wade into auditory perception here – an area I have not researched – other than to note the apparent similarity of processing in that we hold a model of auditory information in our heads (e.g. of a song) and we use that to recognize and appreciate the song the next time we hear it.

The importance of this for my narrative is that the human perceptual system appears to work in a consistent way across the multiple senses. On the one hand, sensory information is processed into manageable chunks of information and this information is then combined with a model of the world that we hold in our heads, which helps us to perceive the sensory information and then take action using that perception.

11. Let's now move our focus to the way we hold information about the world in our heads. It is clear that we hold sensory information along with other information about the world in our mental models. For example, we know that [Usain Bolt](#) is an athlete, that he won the Olympic Gold several times, and we know what he looks like, and what he sounds like. And of course, we know that visual signature of his - who could forget that!



The human brain is clearly very competent at holding a diverse bucket of information about things we are interested in. In the example above, most of us who have watched Usain Bolt in awe hold a large amount of information about him. If we see a photograph of him from the

year 2010, or from the year 2030 (brought back from the future somehow now in 2020), we would instantly recognize the difference. Consider how much detail needs to be stored in our heads to be able to see the subtle differences of aging and to be able to tell that it is Usain Bolt and not someone else.

Consider further the amount of information we hold in our heads of our family and friends, on people we work with, and what we saw on Television etc. It is vast.

12. Nobody will dispute that we hold a large amount of information in our heads. Visual information about people we know are connected in our heads with other sensory information (e.g. a person's voice) along with a large trove of information about that person (what clothes he wears, who his friends are, what work he does, what music he likes, where he lives etc). We have a rich model of the world inside our heads.

Let me then define consciousness as the ability to acquire information about the world, hold it in a model inside our heads, perceive changes to the environment using that model, act using current sensory data combined with that model, and update the model using the latest analysis.

So for example the classic knee jerk reaction ([Patellar reflex](#)) is an example of a non-conscious reaction in humans, while much of day to day human activity is conscious activity. This gives us an initial yardstick for determining whether animals are conscious for instance. Which end of that scale do various animal activities fit? Is it a simple reaction that does not rely on an internal learned model of the world?

We are likely to find a continuum of capability between the simplest animals, to more complex animals, to primates and to humans.

13. Consider now the model we have in our heads of the world around us. It contains information about our friends and family, about the people we work with, about people we see on television, about objects we see as we travel around the world. There is one quite important thing we haven't yet mentioned... that is *us*! i.e. the observer. Our model clearly must have a model of ourselves too. We know we exist, we can see ourselves, that we have hands and legs and a stomach. And of course, we can see ourselves even better in a mirror, or if someone takes a photo or a video of us.

So, the model of our world is not quite complete until we add ourselves into that model. Not only do we have a physical model of ourselves, we also have a model of our own consciousness. We know we exist, that we are conscious, that we have fears and desires,

that we are alive, and we know we can act based on our fears and desires. We have a detailed model of ourselves inside the model of our world! “Self consciousness” is the consciousness that the self exists. It happens when the model of one’s world includes a model of one’s self.

The sophistication of our world model does not end with the inclusion of the self. We know that people around us have models of themselves inside their heads. And we know that those people have models of us inside their heads. And we know that they know that we know etc...

So, in fact our models are super sophisticated. We can run very complex what-if scenarios of how we might respond to an event and how other people might respond to that same event. For example, if one asks a person one fancies whether they’d like to go to diner this evening, how will they react? How many of those scenarios have we all run in our heads in a lifetime? Or if one asks one’s boss for a pay rise, how will she respond? Or if one is playing a game of cricket with one’s friends and you hit the ball more aggressively than you ought to – what might happen next? A good outcome or a bad outcome? Your forecast of the situation changes second by second as the ball flies through the air.

That then is the pinnacle of human consciousness – we can have a very complex model of the world that includes models of other people and how they think, and then do what-if forecasts of how the world around us might react to something we might choose to do.

Consider how incredibly sophisticated that is. Most animals cannot do that. No computer can do that today.

14. Theories are useful if we can imagine new scenarios using them, and we can verify forecasts we arrive at by setting up experiments. So, let’s try some.

Are dogs conscious?

Well dogs are not quite as good at “consciousness” as we are. They do not have a model of New York City in their heads, and do not have an urge to get on a plane to visit it.

On the other hand, dogs clearly have a pretty good model of their immediate environment. They know each of the occupants of their human family very well and they can tell the postman and strangers apart from the family.

Do they have a model of the conscious-being that is inside each human family member? I do not own a dog, so I'll let dog people answer that one, and I expect a debate. It seems to me that dogs have a clear understanding of what "walkies" means for instance, and it invokes a complex expectation of fun when they hear that word. It also seems to me that they can forecast forward about sticks being thrown, about retrieving them, and about sticks being pretended to be thrown, and the fact that "yes I have seen that trick before" etc.

Do dogs have a model of themselves inside their heads? For example – does a dog know how big its own body is? If they see a gap in the fence of different widths, and they can see a ball on the other side of the fence through that gap, does their behaviour depend on the ratio of gap to body width? So, if the gap is just big enough, will they try to wriggle their way through it? If the gap is clearly too small for them to get through, will they try nevertheless? I don't know – I am not a dog person. But no doubt an experiment worth doing. I would like to know whether dogs have an understanding of their physical dimensions. Do they know how far their paw can reach? Do they know that the paw belongs to them? It seems to me that we can setup experiments to verify. We know for example that [dogs do not recognize themselves](#) in a mirror. But that could be because visual perception is less important for dogs when compared with their sense of smell, and they therefore spend little time wondering how they look (unlike us!). There is [some evidence](#) that they recognize their own smell.

Now a bigger question. Do dogs have a model of their own consciousness inside their heads? Are they conscious of their "self"? Do they know that they have desires and that others have desires? Do they have plans? Can they "plot" to get a family member to take them walkies? On balance I think maybe not. But remember, I know very little about dogs – but I do love dogs in case you are wondering – especially other people's dogs...

15. "What is it like to be a bat?". It is difficult to write about consciousness without addressing bats, and Thomas Nagel's [famous article](#) on the subject. Nagel asserted: "But fundamentally an organism has conscious mental states if and only if there is something that it is to be that organism". This highlights Nagel's underlying assumption that there must be a common "subjectivity" that is shared by all bats, that is inaccessible to us humans. I do not doubt that bat subjectivity is inaccessible to humans – but I do not agree that there is a common subjectivity that bats share, other than what is endowed upon them via their shared physical perceptual apparatus. We can however all agree that bat-subjectivity is likely to be more similar between bats than any shared subjectivity we may have between bats and humans. Nobody will argue with that.

Let me address this by focussing on the human experience of subjectivity first. Here's a thought: I do not even know what it is like to be my wife. Even though my wife and I have shared a life together since we met in 1981, when we were 23 and 20, and we have grown

older sharing a house, a family, friends, relatives and holidays, we do not have identical subjective experiences. Of course, we share a love of food, and music, and places to visit, but at the same time we have another list of food, music and places that we do not both equally enjoy. How is that possible? How is it that my subjective experience is not identical to hers? After all the years together?

As we know this is true not just of married couples – it is true of any two people, true even of [identical twins](#) that have grown up together. But that is not so hard to explain. Each person's consciousness exists within the model they have of their world. This consciousness starts diverging for identical twins from the day they are born. They will experience the world subtly differently from day one, and then more and more so as their two models of the world diverge.

Over many years, layer upon layer of experience adjusts and extends the model we have of the world within our heads. We then use that model to recognize new incoming sensory information. There are positive and negative feedback loops inside our heads where incoming perceptions are seen through the lens of past experiences which then create new memories that are fortified (or tainted) by what has happened in the past.

16. Let's address human subjective consciousness and qualia. [Wikipedia defines Qualia](#) as "... the perceived sensation of pain of a headache, the taste of wine, as well as the redness of an evening sky." It surprises me that I could find little written about our sense of taste though in this context. Philosophers often ask about the subjective sensation we have when we see the colour red. But what about food? It seems likely that the human perceptual apparatus for taste is no different than for vision or sound. That is, I assume there are taste sensors connected to short term memory and longer-term memory. I further assume that we extract salient features from the sensory data and combine that with the historical model of food held in our heads to enjoy our current experience of food. I assume these facts as I did not find much research on the subject. Now let's then ask the big question about our subjective experience of food. Do humans have an identical shared subjective experience of food in the way that some philosopher's assume we share a sense of the colour red? Absolutely not. It amazes me how some Westerners can dislike non-Western food, how some Asians can dislike non-Asian food, and so on. There appears to be little innate-subjectivity of food shared by humans.

I travel a lot as part of my other life as CEO of KAL. Let's take the example of six countries with outstanding cuisine that I love – six countries that are all "foreign" to me. China, France, India, Italy, USA and Japan. I exclude from this list my own country of birth Sri Lanka and my adopted country Scotland – because of course my food-model has had years of getting used to that food, and of course I love that food too. You would not be surprised if I said I love Sri Lankan food, or that I love Scottish food at the best Scottish restaurants.

When I visit the six countries I listed, I am a foreigner. But I have visited each of the six countries more than 50 times each, and I love the food. I insist on eating locally when I travel. The point I want to make is this: it amazes me how people can find the best food in other countries unpalatable. Some examples:

- I was at a sumptuous western buffet at a leading hotel in Europe for dinner with one of my Indian work colleagues. He ate nothing, while I made a pig of myself. He went to bed that night on an empty stomach as there was nothing he wanted to eat, nor try - however much I coaxed him.
- I once entertained a delegation of Chinese customers at a US steakhouse. They were very keen to check out the famous Texan steaks they had all read about. And yes, we ordered the best steaks as you can imagine. And yes, at the end of the evening the steaks had been hen-pecked and my guests were starving. (We had to order noodles when we went back to the hotel).
- It amazes me how some western people I know refuse to try a beautiful runny Gorgonzola, or a camembert because it has “white stuff on it”, and even if they do try it after much persuasion, it comes straight out of the mouth.
- It amazes me that some French people cannot countenance an English cooked breakfast with bacon and eggs in the morning (only croissants are ok in the morning), and some westerners cannot imagine having an Indian dosa with sambar for breakfast (“oh no, not curry in the morning”).
- It amazes me that some Italians will say that *only* Italian food is great, and some French will say the same, and some Indians will say that non-Indian food is tasteless, and some Japanese will say that western food is full of milk and hard to eat, and Chinese colleagues will go hunting for Chinese food when they travel abroad, and some Americans eat only hamburgers when abroad, and some Brits go looking for what they eat at home when they go on holiday. Spanish beach resorts have areas that are full of English eateries just for Brits on holiday.

Of course, many people are not like that. But you get my point. It is clear that human subjective consciousness of food is rooted in what we have already eaten and enjoyed in the past, and what somebody such as our parents coaxed us to try. Not many people do what I do – eat locally despite initial impressions which may not at first be positive; train your taste perception over many visits to get used to the local cuisine; and learn to enjoy Chinese food like a Chinese person. (Try to avoid the bland Chinese food served in the West – unless you can get them to serve from the secret menu behind the counter that they reserve for Chinese customers only).

Let’s now address the classic philosophical question about how hard it is to explain what colours are to a person who had been blind from birth. It seems to me it can be equally hard to explain to an Indian person who refuses to try French food how wonderful and subtle French cuisine can be (and vice versa). The point I want to make then is that just as much as the lack of sensory equipment can be a barrier to understanding from another’s point of view, so can the brain-model and a lifetime of perceptions built-up inside that brain-model over time.

So, what are the lessons from this digression into food?

- There isn't a shared human subjective consciousness of Gorgonzola, or sashimi, or ratatouille or curry.
- The food we like is hugely influenced by the food we have eaten all our lives and that has conditioned the food-model inside our heads.
- At the end of the day food-consciousness is a very personal experience. No two people are the same.
- A person who hates Gorgonzola has as hard a time to understand how someone else can love it, just as much as we find it hard to understand why Nagel's bat loves insects. Would a bat brought up on Gorgonzola from day one learn to love it? Probably. I feel an experiment brewing.

17. My conclusion on human subjective consciousness is this. There isn't a shared magic that binds together the consciousness of a species. Nagel asserts "But fundamentally an organism has conscious mental states if and only if there is something that it is to be that organism—something it is like for the organism". I believe this is wrong. There is no special magic that binds together the conscious mental states of an organism – other than those forced upon it through shared sensory equipment and shared lives. To the extent that an organism has a model of the world inside its head, that model is a creation of the sensory equipment and its past life up to the current point in time.

I also believe that Nagel's "if and only if" is especially incorrect. There is no reason that an organism *must have* a shared subjective experience with other organisms of its own species in order for that organism to have conscious mental states. Indeed, a bat that is born with no sonar equipment due to a birth defect is still a bat – but surely with a very different subjective experience from its fellow bats – and surely with its own mental states nevertheless.

Consciousness is a very personal experience. There clearly are shared experiences that we can agree about (my wife and I love certain Amarone wines for example), while we can equally disagree about other things (I love kiwi fruit, but she will not touch it). How is that possible if we are both human and we have shared so many years together? What happened to the kiwi-fruit qualia shared by the human species? Apparently kiwi qualia is not built into all humans.

If you are still not with me on this qualia thing – let me try one more time. Think about a rock band or a famous musical outfit that you *do not* care much for. Now search for a video of them on Youtube, playing in front of a huge crowd – perhaps in a stadium with tens of thousands of people. Now watch the audience get themselves into a frenzy as they listen to a piece *they* love. Now ask the question, what happened between you and them on musical qualia? Where is the shared qualia when it comes to music, between people of the human species? In fact, there probably is very little shared musical qualia. Think about the

difference between Western music and Indian or Chinese traditional music for example. And what about the folk in the Amazon that have not joined the rest of us? What music do they like?

18. That brings us to the so called “hard problem” of consciousness. David Chalmers [wrote in 1995](#) “Why is it that when our cognitive systems engage in visual and auditory information-processing, we have visual or auditory experience: the quality of deep blue, the sensation of middle C? How can we explain why there is something it is like to entertain a mental image, or to experience an emotion?”. Chalmers’ underlying assumption is that there is a magical experience of middle C that is shared by humans.

As we have already seen it is unlikely that there is a quality of middle C that is shared amongst humans. Chopin’s appreciation of middle C must surely have been a universe apart from my own – me being a musical dummy.

To me, the hard problem does not exist. Chalmers begins with an underlying assumption, as does Nagel, that humans share a magical subjective appreciation of sensations such as middle C or the colour red. But as we have seen from my foray into Sashimi and Camembert, that we certainly do not when it comes to food. And I can absolutely guarantee you that I as a human being with no musical talent whatsoever, have no magical subjective experience that I share with Chopin or Paul McCartney. None. Zero.

When I listen to a musical piece by Chopin – the likelihood is that that subjective experience is unique to me, and different from the experience other people around me might have, and very different from the experience Sir Paul might have – whether he likes Chopin or not.

There is no shared subjective experience other than what the physical equipment of my sensory equipment endows upon me, coupled with my (rather minimal) knowledge and brain-model of Chopin’s music.

Chalmers categorises past approaches to the “hard problem” thus:

1. “Some researchers are explicit that the problem of experience is too difficult for now, and perhaps even outside the domain of science altogether.”
2. “The second choice is to take a harder line and deny the phenomenon. ... According to this line, once we have explained the functions such as accessibility, reportability, and the like, there is no further phenomenon called “experience” to explain”
3. Plus three more with diminishing relevance to the problem at hand.

Chalmers dismisses these approaches and proposes his “extra ingredient”.

This then is the crux of the problem. We as human beings are determined that there is an extra ingredient to our conscious experience. Some Ether perhaps, or dark matter, or quantum mechanical activity that makes humans truly unique. Just as much as we were once very sure that the sun revolved around the earth, humans are constantly searching for evidence of our super humanness. Consciousness appears to be the last bastion that science has not conquered, and therefore a breeding ground for outlandish theories such as those on quantum mechanical influences.

Having said that, Chalmers' theory is not one of them, but is ultimately unsatisfying. His addition of an extra ingredient in the form of "information", does not seem to progress us by much.

To me, Qualia are simply experiences that we have learned over a lifetime and that have become part of us. They are held in our brain model of the world. We add to them each day when we perceive our world using the brain-model, and then add back to that brain-model as we go.

19. We cannot wrap-up on the philosophical view of consciousness without talking about Searle and Dennett.

I have to admit that I did not devote a lot of time to understand what Searle and Dennett have to say. I apologise therefore if I have not done them justice.

[Searle](#) it seems to me insists that there is a magical property to consciousness that has not been explained by anybody so far, and it seems he does not expect that it ever will be either.

[Dennett](#) on the other hand is equally determined that consciousness is nothing more than what the physical brain can do and there is no need to assume that any kind of magic is needed at all.

20. And of course Richard Gregory's work:

[Gregory](#) argues that there are top-down processes (i.e. in my words, where the brain-model of the world is trying to work out what is being seen), and bottom-up processes (i.e. where sensory data is being processed). Examples of optical illusions such as the [Necker Cube](#) are

clear and easily observable examples of the way the brain-model can flip the perceptive process. Gregory rightly focusses on a series of illusions.

Gregory says:

“In the account given here, perception depends very largely on knowledge (specific 'top-down' and general 'sideways' rules), derived from past experience of the individual and from ancestral, sometimes even prehuman experience. So perceptions are largely based on the past, but recognizing the present is essential for survival in the here and now.

The present moment must not be confused with the past, or with imagination, i.e. as indeed one appreciates when crossing a busy road. So, although knowledge from the past is so important, it must not obtrude into the present. “

So right.

21. Idea: How is the brain-model and qualia created? At birth it would seem reasonable to assume that we have a basic set of qualia we are born with. Like hunger. Any gene that interfered with that would have been rooted out millions of years ago as that baby would not have survived to pass on its genes. Starting from a basic set of qualia, and equipped with sensory hardware, a baby begins to make sense of its world day by day. It notices that hunger pangs followed by instinctive crying results in a huge blob (that it later learns to call mum) to come over and sort things out.

I suggest that qualia are built-up layer by layer over a lifetime. I also suggest that all of us create our own set of qualia – with significant similarities of course – as we have similar sensory equipment and we live similar lives. The qualia-rich brain-model evolves as we grow-up, it is fortified as we learn to speak and go to school. This is further reinforced when we begin to read and start to watch television. That allows us to absorb a massive amount of knowledge about the wider world outside of our immediate environment.

It is clear to me that there aren't qualia that are automatically identical between humans *just because we are members of a single species*, other than a basic set we inherit at birth. For instance, it is possible that we see the colour red the same way perhaps, but we certainly do not all see the colour yellow the same way... my dad and I saw it differently for instance, as he is red-green colour blind while I am not. (The colour yellow requires both red and green receptors to work to be able to discern).

My foray into food earlier demonstrates that food qualia are not shared amongst us as a human species. Much of it is learned - the love of blue cheese does not arrive automatically along with the initial box of genes we receive at birth.

I theorise that we humans start with a basic set of qualia on day one, we do the best we can with that minimal set to make sense of the world, we add to our brain-model slowly at first, but faster as our brains develop and we become hungry for knowledge, encouraged and influenced by people around us.

As I mentioned earlier, even [identical twins](#) with identical genes and an identical home environment grow up to have different qualia.

The brain-model is pivotal to how we see the world. Humans might perceive visual stimuli in similar ways, most likely due to the high quality of our visual sensors and the importance of getting it right for day to day survival. It would be a bad idea to mistake a bus coming down the road for a Christmas cake – people who do that do not survive to tell the tale. But once we move away from visual stimuli, our ways of seeing the world differ very significantly. We do not enjoy music or food the same way at all. Engineers, doctors and lawyers perceive their work tasks very differently – a hardware engineer that is able to fix a [PCB](#) would be no good at all in trying to help out with open-heart surgery.

Let's touch on politics for a moment. We see massive polarisation around the world today. People perceive political events through their own set of political qualia. Unfortunately, the perception process often results in what electronic engineers call a "[positive-feedback loop](#)", where a certain way of seeing the world is further enhanced by all incoming fact – irrespective of whether the facts agree with the internal brain-model or not! The brain-model overrides reality and sees it the way it wants to see it – "huh I knew it – I was right all the time"! Just like the Necker cube illusion - you see it the way you want to see it. OK enough said about politics.

This analysis also has an important perspective on [traumatic events](#) people might experience. As we know people can suffer from psychological trauma for years. The general narrative is that the original memory of the traumatic event keeps haunting the person. I think this is not the whole story. True, the traumatic event creates a virulent memory in the person's brain. And yes, all incoming perceptions are impacted by the affected brain-model. However, what happens next is as important. If the person continues without help, new memories are laid down day in and day out, layer by layer, *influenced* by that original trauma. This can create a positive feedback loop that exacerbates the traumatic memory through new memories that *reinforce* the initial trauma. People in that situation need help. Medication can help as we know. The most important thing is that the brain's built-in mechanism that slowly forgets older memories should be allowed to take its course. It is important to build a new brain-model that narrates away from the initial trauma.

22. As we come to the end of this article, let's celebrate what human consciousness is. Consciousness is about creating a brain-model of the world around us, about using it to perceive the world, about using it to forecast what will happen next, about using it to understand other people that matter to us, about using it to understand one's inner self, about using it to understand how that inner self might interact with others that have their own inner self. It is about acting in that world and learning from those actions. It is about improving the brain model bit by bit to end-up with better outcomes for one's self.

Consider for a moment some examples of how sophisticated our human consciousness is:

- This is a [favourite video](#) of mine when Jonah Lomu scores a try against England in a world cup rugby match in 1995. From the eighteenth second onwards when Lomu picks up the ball, stop and start the video every second for the next ten seconds. Think about what is going on in Lomu's head and what is going on in the English defender's heads at each moment. Think about their individual brain-models. Lomu was already famous – the English defenders knew what he could do and what to expect - they had studied him carefully before the match. Consider also that each player must stay within the rules of the game at each moment, not to mention the laws of physics and gravity, control his own body and not fall in a heap, and push each muscle to its maximum capability. Watch the situation as the game evolves second by second. Each player assessing his ability to stop Lomu, while Lomu considers how he counters that. At the 22nd second Lomu thinks – “am I going to be able to keep my balance – and surely Mike Catt the final defender is well positioned to stop me given I just lost my footing – I am up to Mike Catt now in the 23rd second already – oh and I run through him – nobody expected that – not even me”! (My apologies if rugby is not your thing – hope you enjoyed the ten seconds of rugby video). I can tell you I set down some unique qualia at that moment in 1995 as I watched it live on TV. I still feel it each time I watch that video.
- Let's consider a competitive situation that evolves at a different pace. Consider a contract negotiation between a customer and a vendor. The two teams start from very different perspectives. The buyer wishes to buy something they need to have, the seller presumably specialises in producing the thing that is wanted by the buyer. The buyer is probably not expert at buying this specific thing – the purchasing people would normally buy a variety of things. The seller on the other hand is likely to be specialised in selling that product – they do that day in and day out. The buyer wishes to buy at the best possible price but also wishes to secure contractual terms that provide long-term security for his purchase. The two teams know that this needs to be a “win-win” situation, but they start from opposite ends on price. Both teams need to understand the deal from the other's perspective in order to finally ink the deal. Think about how the brain-model of each team member evolves over the period of that negotiation.
- Finally imagine one more scenario. A family walking down the street – mum and dad with two young children. It's a lovely sunny day. Each person has his own brain-model and his own qualia as they walk happily down the street. Now in your mind's eye stop and start the video every second of how the four individuals react when the youngest one spots an ice cream seller. Oh yes you know how that video runs. “Mummy

mummy” it begins. Now imagine how that video might run if the family had only just finished eating an ice cream 10 mins ago. Isn’t it amazing what scenarios you can run in your brain-model? And I am sure you can see it from all four perspectives – unless you are only 4 years old!

23. Finally a few words on mind-body dualism and René Descartes. A debate has raged since Descartes said “I think therefore I am” in 1637 in his [“Discourse on the Method...”](#). How are the mind and the body related? To me as a Computer Scientist this has always seemed easy to answer. The Mind = software and the Body = hardware. OK I get that it is not that easy. Especially so as the human brain probably does not separate hardware, software and data as we would when we design our systems. It is perhaps the case that when the brain stores a new memory, that [changes the neuronal hardware](#) very subtly. But that is true with our computer systems too, depending on how deeply you dig into the mechanism of storage. Although we might say “no way does the actual hardware change” when we modify some data on a computer – but indeed the hardware *does change* very subtly at the level of the physics of the RAM, or the SSD, or the hard disk being used.

This paper is primarily about a computational view of perception and consciousness. Therefore, the way in which the brain might implement the mind/body duality is less interesting here than how we might do so in a computer. Consider that when we built cars, we did not replicate legs. When we built aircraft, we did not replicate flapping wings. Our computers today do not for the most part have a neural-network architecture – and I do not believe we need to do so.

Computers have hardware that we consider is fixed, and software that we can change as we like - a computer can be deployed to perform accounting activity or research into visual perception, without changing the core hardware. We further categorise software into a “fixed” program and “changeable” data. It is considered bad programming practice to allow the program to modify itself. But that is not a fundamental rule – just good practice that allows us humans to understand and manage what is going on inside our systems.

Should we try to mimic the human brain? I don’t think so. It does not surprise me that when we try to understand the way the brain works we find it so hard to do that using the technologies we have today such as electrodes and magnetic detectors. Imagine trying to work out what a computer program is doing using electrodes on the motherboard - you would not get very far.

So, can we build a computer system using current technology that can be as successful as a human at visual perception? I think we can. My ideas on this are in point #9 above. But there is a point “#9f”. Without generic knowledge of the world it cannot compete with human vision.

So, can we build a computer system that can be conscious like a human? I have to say that my answer is a cautious yes. See #21 above. And will that system have qualia like a human? The answer must be that it will have its own qualia. But can we try to make it as close as possible to ours? As human designers we surely will.

24. In wrapping up this paper then, and at the risk of upsetting Mr Descartes, I will summarise my theory of consciousness as follows: "I model, therefore I am".