

*Gresham Lecture, Wednesday 16 February 2011*

**Why we see what we do**

Professor William Ayliffe

Thank you very much for coming along this evening to the final part of the series of lectures that I have done trying to explain how we see things. We started off, some time ago, looking at the anatomy of the visual system – we will reiterate some of those things where they are relevant, we then looked at where the so-called visual system goes wrong with illusions. I tried to explain that, in fact, it is not going wrong – it is just the way we see. Now, the combination of both of those leads us to the topic of why we see what we do.

Matthew Luckiesh said: “Seeing is deceiving”.

The rather eccentric son of a Swedenborgian, and brother of Henry and Alice, William James, the famous brother who was an author, was actually a clinical psychologist. He said about 100 years ago that “Whilst part of what we perceive comes through our sense from the object before us, another part (and it may be the larger part) always comes out of our own mind.” It is really astonishing because we all think that vision comes from our eyes when it probably does not. We are probably making it all up. Not only are we making it up, it has probably been made up before us, generations before, and I am going to try and show you why we think this is the case.

All we know about a particular object is that it has a size and it has a direction. The shape of the image does not give us any information about this. We do not know how far away it is. The shape could represent any of an infinite number of sizes and distances. For example, the image of someone as they walk further away from you necessarily becomes smaller and smaller, as you know if ever you have taken a photograph. Actually, in our eyes and brains, that person does not become smaller.

Let us assume that the human visual system has evolved over millennia from very primitive, call them, ‘shark detectors’, assuming that we all came out of the ocean. This brain is programmed for interpreting moving, three-dimensional objects, and it has to interpret in an environment where these objects are distorted by perspective. So, the retinal image, whether it comes from a real three-dimensional object, or whether it only comes from a flat two-dimensional object, such as a painting or a photograph or a cinema film, is flat and uncoloured, but our perception of this image is very rich. It is in three-dimensions, it is coloured, it is textured, it moves, it is vivid, and furthermore, it occurs in a cluttered scene which is also three-dimensional. Even data from two-dimensional images can be transformed into these three-dimensional perceptions.

The retinal image itself, as I mentioned, is ambiguous. It could represent an infinite variety of possible objects, not only shapes and sizes but also luminance and reflections. We do not know any of these things from the real world. All we have is a retinal image. Perceptions are also influenced by previous experience, and I showed you, in my visual illusion lecture, even sexuality can change how you see things that are put in front of you. Finally, it may be influenced by what our ancestors saw, and these ancestors may not have even been humans.

The great unsolved mystery of the mind, which still remains today, is how we see. Vision appears to be effortless and instantaneous. We open our eyes in the morning and everything seems to work – we just take it for granted. Seeing begins with an image, but it ends with perception, and perception and an image are completely different things. Perception is not a copy of the retinal image. Actually, it’s constructed in the brain from a whole heap of very conflicting data. All of these data points end up giving us exactly the same image. They are all at different distances and different angulations yet we get the same retinal image from them, and this problem involves massive computing power. Roughly half the cells in the brain receive visual input, so half of the human brain, which is a remarkable complex series of computers, is involved with vision. However, it is also conscious. This information from the retinal image is processed by the brain to generate a percept, and this perception, does not match the physical properties of what caused it originally so perception and so-called reality are two different things.

When I told you that you are making it all up it sounded very stupid but I can show you that it is actually true.

When we take a picture we take a photograph and we develop it. It is designed to be looked at by the human eye, and then it forms another image, which is the retinal image. The retinal image is not a camera image as it is not been designed to be looked at by another human eye. Remember the talk on the murder of Annie Chapman, when they did not take photographs and the argument in court was that they could have actually imaged Jack the Ripper on the back of her eye if only they had taken photographs, which they did not, and the police were actually accused of being negligent.

If we did have a retinal image that was looked at by another eye, we would need another internal eye to look at this and then pass the information on to the higher centres. This version of how the brain works was pretty current when I was at school, and pretty much what we thought for several years afterwards. Remember that the image on the retina is only a series of brightness numbers. It is a bunch of pixels, let us say, from one to 256. If you have done that on a little tiny square, you end up with an infinite number of numbers – it is a very complex thing to do. This image is processed by the retina using highly specialised analogue computers, the so-called centre surround receptive fields. This enhances the light and dark borders, and extracts relative luminosity values. It does not tell us how bright anything is, only whether it is brighter than the thing next to it, or brighter than its background.

Image formation has been known about for a long time, as we can see from this glass ball focusing light from a source which comes from the Opus Majus of Roger Bacon. The first person to realise that the eye did this was Johannes Kepler. He worked for Tycho Brahe, whom he subsequently succeeded as Imperial Mathematician, and wrote a pamphlet called Vitellionem paralipomena, a supplement to Witelo. Witelo was Dominican and actually stole Bacon’s work and published it himself, as we showed in one of my earlier lectures.

Kepler showed how cones of light from an image, from an object, form an image. This image is necessarily upside down in the back of the retina. This did not surprise him – they knew about pinhole cameras which also invert the image. He also realised that this was not something that was looked at by the brain. He said that the perception of this image “I leave to the natural philosophers to argue about – maybe it’s in the hollows of the brain due to the activity of the soul.”

There are problems with retinal images, and there are also problems with images that come from cameras. All optical devices, by necessity, have a number of imperfections in them. The lens is based on a spherical configuration, so it causes spherical aberration. You can make aspheric lenses, as we saw from those Viking lenses in the lecture on spectacles, which do remarkable feats. The projector lens that we are using here has an aspheric lens in it which tries to get rid of the aberrations, but it is possibly not as effective as it might be.

There is also chromatic aberration. This happens because a lens focuses blue light in a different place from yellow light, into a different place from red light. You also have optical aberrations. The Hubble Telescope had optical aberrations when they first sent it up and had to be corrected mathematically. The human eye has optical aberrations, as seen by these aberration diagrams, where red is showing the focus in front of the retina, blue behind, and these images are not perfectly focused. This is normal.

This is what we normally see. This is Van Gogh’s painting at Arles and the Bend in the River. Actually, if you go there, be prepared to be disappointed because the actual polar star is not there, it is round the other bend in the river; I spent several nights being a bit disappointed, not working this out. He is showing a blur there, and it is not because Van Gogh could not see. He was painting the normal spherical aberrations of the human eye, which increase as the pupil enlarges. The wider the aperture of the camera, the more aberrations you will get.

When the eye is focused for mid-spectrum colours, such as green, the blue light is blurred and cannot contribute, so it causes blur. Some people have thought that the back of the eye in the centre has yellow pigment and the macula – sometimes called the macula lutaea, which is Latin for yellow - because it absorbs the blue light and stops the blue cones messing about with the image quality. Furthermore, there are very few blue cones in the centre of your eye; they are more distributed to the mid-periphery. It is a nice theory, but, unfortunately, the facts get in the way of all good theories, and it turns out that monochromatic blur is many times worse.

This is not to say that the retina does not form an image. It does, and as we saw, this is the optogram that was taken from a rabbit retina immediately after sacrifice by Willy Kuhne. He fixed the retina and has the image here of the bar of the rabbit’s cage and the window of his laboratory. This created great excitement at the time and this is why people thought you could get images of murderers on the back of victims’ eyes. Furthermore, Willy Kuhne did not help it because he claimed to have obtained an image from the retina of an executed criminal in 1880. However, the point is, is you have to fix it, because this visual pigment decays in seconds after death, and after 60 seconds, it has completely gone. Very few ophthalmologists have ever seen the beautiful colour of visual purple when we examine eyes, partly because we bleach it when we look in with bright lights, but the main reason is, even in dissection, by the post-mortem, it has already gone.

We mentioned these analogue computers called receptive fields, and I think I should briefly remind us what they do. They were discovered by putting an electrode into the retina and recording from ganglion cells. In the complete dark, ganglion cells fire – ‘tick, tick, tick, tick, tick’ – which is quite interesting, and if you shine a bright light on the eye, the ganglion cells go ‘tick, tick, tick, tick’, but they do not change. That really freaked people out. That was really unusual, because most people think that when the eye sees light there is some response. Actually, if you shine a small spot of light on certain cells, these on-cells, you get a burst of firing and if you actually shine it on these type of cells, you get ‘tick-tick, tick-tick-tick-tick-tick, so they stop firing. These are called off-cells and these are called on-cells. If you move the spot off to the side, the on-cell immediately stops firing, which is interesting. Here we have an analogue computer that compares. It subtracts the light falling on one part of its receptive field from the other and shows that light and on and off-cells are equally distributed in the retina because dark things occur in nature just as much as bright things, rather like the print on the back of this slide here is dark against light - we would be using off-ganglion cells to represent that.

At the end with this processing we tend to highlight edges of objects. We tend to see them enhanced more brightly, and the dark areas become darker so this helps us process the retinal image.

After we leave the eye, remember we go through a crossing. The interesting thing in this ancient Arabic manuscript is you can still see that the seat of vision is believed to be in the centre of the eye in the lens. By the time that Chevalier John Taylor drew this diagram, it was realised that the lens was a focusing mechanism. However, Taylor took further Newton’s discovery that the vision was crossed and to show it is half-crossed – it is a hemidecussation of the image. This means that everything from this half of the field of these two eyes goes onto this eye, and goes to this part of the brain, and everything that is on this half goes to this bit and goes to this half of the brain.

The digital output from this retina has now been sent on, after it has crossed, to this organ, the thalamus which is in the centre of the brain. It is about the size of a walnut, and think of it as the switchboard of the brain. All sensory input, whether it is from fingers, toes, hearing or sight, uses the thalamus for a relay. 85% of the input to the thalamus, however, is not from sensory organs at all. It comes from centres that are aware: attention, awareness, motor functions and sleep. Also, it gets a lot of input from the primary visual cortex, so the vision centres come back down, and nobody knows what they do, but it must have something to do with top-down processing, telling us something about what we already know before we send it up. The remainder of it is from the eye.

There is another little nucleus here, which we are going to come to, called the superior colliculus. This represents an archaic nucleus present in primitive forms of life called the tectum because it is the roof of the original brain. This is concerned with eye movements and head and neck movements. We map the image and there is a point to point relationship, as Kepler showed. Everything on this half is here, crosses over and goes to this lateral geniculate nucleus in the thalamus, while everything on the green half goes over to the other one. There is a duplication of these parvo-cells and the M-cells although nobody knows why. It could have been an accident; it could have a function. There are also branches which come off to the superior colliculus as well, representing that archaic pathway we mentioned.

What is the purpose of the retinal image? It is to compare the number of photons absorbed by different photo receptors. You will get a larger and better quality image from larger eyes, which is why nocturnal animals and birds have large eyes. The image is converted into a map, as we showed you. This is a primitive bug detector on the left-hand side of a frog, and this little juicy tasty morsel flies into place and goes onto the receptive field, which immediately fires. This one does not. This fires and it goes onto the roof, where it is mapped very accurately. From this map, immediately beneath is a motor map which is connected point to point. Automatically, this goes to the spinal cord, the muscles on the other side so and the neck and the eyes move towards the prey, the tongue comes out, and that is the last thing that that insect ever sees.

It is quite interesting because Jacques Vaucansion invented these wonderful automatons which actually influenced how Decarte thought the body and the brain worked, because they were completely automatic. We have a completely automatic system here. The frog does not have to think. Over here, tongue out, back in, dinner!

Here is a real image. The projection seems to be dark here, unfortunately, but you might just see a little white clot here, which has blocked off the retina in this part and starved it of its oxygen supply and so it has died. The visual field effect we get is down below. This is the same patient. So we’ve got an inversion. Now, the inversion actually does not really matter, and it did not bother Australians for the first 200 years of their existence, until they then invented this map because they suddenly realised they could not navigate around unless they had the map of the world showing it from an Australian perspective. I think there are maps of South London that can help spouses navigate when you go to Croydon and Brighton – but I have not seen any yet.

We also know that different parts of the brain do different jobs, and this was introduced by Albertus Magnus. Here is this beautiful fresco of him in Treviso. This cycle shows the earliest images of anyone wearing glasses or any magnifying aids for vision. This is Albertus Magnus, who is further round in the series.

After the thalamus, this processed information goes to the occipital lobe of the brain, which is at the back. The scene is then mapped again and further processing takes place. This edited information is then transmitted across the whole surface of the brain, the whole back half, both the temporal and parietal lobes, and goes to specific areas that do specific tasks - what we call the higher centres.

Some of you will know this image. It is actually by a former Gresham Professor, and it is the first accurate description of the brain in world literature. It was done by a sometime architect called Christopher Wren, who was very familiar with anatomy and built a small building just down the end of this street, as some of you saw when you came up from the tube station. He also made models of muscles and models of eyes, and he did a lot of his work by injected India ink and wax into the clotted arteries, and this enabled Thomas Willis, in Oxford, to do the first accurate description of the blood supply of the brain in Cerebri Anatomie. They came up with a really peculiar and interesting theory of how the brain was working. They decided that the corpus striatum was for sensory information, the corpus collosum, which is inside the brain and fibres going across here converted perception and imagination into the hard overlying tissue here, the grey matter, the cortex, which they thought of as the core or the peel of an orange. The main bit of the brain was the juicy bit on the inside, and then the peel was actually not considered to do very much at all, except to store memories. This theory of the brain had widespread influence for several centuries.

Francesco Gennari was dissecting a brain as a medical student, in between going to the pub, as medical students are wont to do. The problem with Gennari was that he did not stop going to the pub and died young of severe alcoholism and lack of money from going to the bookies on the way home from the pub. That is a sad story about a brilliant man, because when he did these frozen sections, you can definitely see that there is a white line here which is not present in this part of the rind or the cortex. This was the first intimation, in fact, that the orange peel is not uniform. There are different bits of the orange peel and they have different structures. Maybe this could mean they were doing different things. He then says: “I do not know the purpose for which this substance was created.”

More detailed analysis, once microscopes had been invented and used, was done by Brodmann. He showed that there are definite distinct variations in the structure of this rind in the brain, and the cells are very different. This picture shows the outside of the left-hand side, and this one shows the inside of the right-hand side of the brain, after we peel the two parts. We find this area here, called 17, which has a highly specialised structure indeed which was what Gennari had found without using the microscope.

We have a temporal lobe here, which comes down to the temple. We have a parietal lobe, which is at the back. We have an occipital lobe, which is at the occipat, and then we have a frontal lobe.

Vision is located in the occipital lobe. Franz Joseph Gall, who, amongst being a complete nutter, invented phrenology, was also a very bright neuro-anatomist, before he went into his daytime job. He noticed at school that boys with phenomenal memory also had bug eyes and he called them “les yeux a fleur de tete”. Therefore, he proposed this was due to an over-development of the frontal lobes. He then studied medicine and he was the first person to dissect crossing fibres that showed that motor function is on the opposite side, a bit like vision. It is quite interesting how we cross everything over, as well as having it upside down. It does not seem to bother us but it is interesting, and he was the first person to show that. He proposed that there are different functions in different areas; the white matter is not the stuff that’s interesting in the brain – that is just the copper wiring. He also proposed that the brain was so folded to save space – if it was not folded like that, we would have brains of enormous size and this would be very inconvenient to get through doors. This was a revolutionary thought, and it was contrary to religion, so he was not very popular.

Bartolmeo Panizza examined some brains of people who had had strokes and who were blind. He found that the back of the brain, the occipital cortex, was crucial for vision. His discovery, of course, was completely ignored, as many were. He was a follower of Gall, and that probably did not help his case because everyone, by now, was convinced that Gall was mad, except for people doing phrenology, and some people still do phrenology. You might see them in antique shops, those heads where it says that you have individual functions in distinct places.

It took some horrible wars to prove this guy right. Head injuries have been known in war for a long time, and these are drawings that were done by the English surgeon after the Peninsular War battles. You can see a very good head injury here, from a bullet wound. His gloves and saw still exist, and he sawed off Uxbridge’s leg at the Battle of Waterloo, saving his life. He was not actually particularly good at saving lives in his earlier career.

On the 17th of September in 1862, the bloodiest battle in American history took place, when General McClellan confronted Lee’s army of Northern Virginia at Sharpsburg. This became known as the Battle of Antietam. The 4th New York Division were on the left, attacking up onto here, and this chap, who had just come from Ireland and joined up the volunteers, got shot in the head with one of these, which was made just down the road in England and shipped out, in large numbers, to the Southern cause. It was a musket and it was very heavy and it made one hell of a hole in his head, and amazingly, he survived. We discussed his story in one of my previous lectures about how he survived, and basically, the surgeons left him alone – if they had fiddled with him, he would have died.

Anyway, he turned up eight years later, where Dr Keen & Thomson examined him in New York. He was complaining that he could not see out of his right eye. However, whisky affected him as usual and his sexual prowess was undiminished. They carefully plotted the visual field, and lo and behold, they found the same that I found when I plotted this visual field in a patient who had had a stroke in the occipital cortex, which is at the back here. This is what he meant when he said, “The vision in my right eye is bad.” What he actually meant was, “The vision on my right-hand side is bad,” and what they showed was that the map of the brain is bisected through the middle, so the map is centred on where you look, on what we call foveal fixation. Fixation is where the map is centred, and this was an amazing and important discovery.

Remember that images are just a pattern of light and dark on the back of the eye, or on a photographic film. A map is a device for transmitting information. It is amazing what sorts of information you can transmit with a map. For example, you can tell people where to go on the tube, and here is an original tube map. You can actually discuss the migration of languages, and this is how the early so-called primitive languages migrated from this area down into Central America.

You could have a spatial map, like the Paris tube map, which is chaotic compared to Harry Beck’s beautiful 1931 wiring diagram, which tells us about connections and how you move from one connection to the other. You can also have non-spatial maps, like Wyld’s map here of civilisations, done in 1815. Morgan colours the countries according to civilisation level: so you have England, which is 5, and you have Canada here, unfortunately, which is only 2 because it contains cannibals and Frenchmen. However, it is elevated above Australia which scores 1 - we are not sure whether this was due to the cannibals or the Frenchmen.

Booth’s Poverty Map just south of Michael Mainelli’s office actually says “vicious, semi-criminal”. Interestingly, walking down this road the other day, I noticed how many of them were hedge fund managers and bankers. However, in a previous time this was Hanbury Street where Mary Chapman was found by her flatmate murdered in 1880, when Jack the Ripper did her in.

Bullets became more sophisticated and guns got a bit faster, so we had more precise injuries. Here is this poor chap who was shot through the brain and ended up with a very specific injury. This enabled the young Japanese ophthalmologist Inouye to start to map, very carefully, what bits of the brain performed different functions and he found something remarkable. He found that most of the visual cortex here is concerned with a tiny bit of central vision. In fact, even more extraordinary is that the central 2.5 degrees, the tiny bit there, is represented by an enormous fraction as well, so there is a massive magnification of vision which is occurring in the brain.

More sophisticated analysis was done at the back of this area, looking into this deep groove, which this picture shows opened up – it is not opened up in life. We are opening up to explain what is going on and a very detailed map was shown, which actually shows what had already been known, but in more detail which is that the bottom represents the top of the vision and the top represents the bottom of the vision in very great detail. You can see that the tiny bit in the centre is represented disproportionately by a large area at the back.

Modern mapping is even more sophisticated and shows that even this was wrong. This is the central area. It is taking up a massive amount of brain computing power to compute what you are seeing in the tiny central bit of your vision. It is called the cortical magnification factor.

Here is a similar thing that occurs in the sensory cortex. It shows that there are bits of our bodies which are slightly more important than others, depending on whether you are a man or a woman. Males and tongues of course are very important and so are hands, and so the brain expands and makes them larger. If we actually drew that as a map, that is what we would get, and we would not get a map that looks like a map if we drew the map of the visual cortex. We would get a map of what the visual cortex is doing. It is using these columns here, which are composed of different cells as super analogue computers. These columns are about 2mm square, and the whole of this cortex is tiled with these 2mm square tiles, each of which is a vertical super computer.

Here they are in more detail. This is the tiling, shown here. It looks like zebra stripes – brown stripes representing left eyes, lighter stripes representing the other eye. They are intimately together, and what happens is that stuff comes in from the left eye, stays in the left eye and goes into Gennari’s white stripe that we saw earlier. Stuff also comes from the right eye and goes into Gennari’s white stripe, but in this column, and these columns at this level, they are separate - monocular. Leaving here, it is then processed and the eyes then come together. This is where the images start coming together and a lot of processing goes on for binocular vision – using two eyes together for depth perception - and for early bits of movement in primates and other mammals, but not in rabbits, who actually have movement detectors elsewhere.

The interesting thing about these is that, on the retina, they had those circular surrounds. These cells do not respond to circles, they respond to lines, and these lines have to move and be tilted very precisely for that particular cell. This cell is orientated to a vertical one, and as soon as you go off vertical, the response starts going down. When you have a vertical line the response is tick-tick-tick-tick, and we inhibit its response when we move to the horizontal. These receptor fields in the periphery are fairly large and they are rather small in the centre, so the size near the foveal is about a quarter of a degree, which is about half the size of the moon subtending an angle on the eye. In the periphery though, they are quite large – they represent one degree.

There are different types of these. There are complex cells, and about 20% of these cells respond to a particular direction of movement. So not only does your bar have to be moving, it has to be moving in a certain direction, which is interesting. These receptor fields are much larger and because you have complex cells, the Americans discovered that you have to have hyper-complex cells. These come in two varieties – they are called end-stopped and then there are different types. These are responding, as you can see, in different ways, so the bar has to be just so. If it is too small, it does not respond; if it is perfect, it responds maximally; and if it is too big, these hyper-complex cells stop responding again. These hyper-complex cells tell us about sizes, widths and lengths of images. They also tell us about borders, how they cross and helps us determine where edges are of curved and squares.

We have moved a long way from looking at different pixels. Here, we are looking at Gabor patches. These patches, if we put them altogether, are telling us about texture. You need a lot less stored information if you are looking at textures of scenes than you do for pixels of scenes, because just a small one millimetre area at 256 x 256 on your camera is going to take up an unbelievable amount of computing power! We would probably need brains the size of this room if we were going to walk around and use that as the way we were going to process the image. Gabor patches are much better, and these actually represent, even artificially, really rather original and realistic looking textures. The visual scene turns out to be full of these textures, and using Gabor patches, we can already start to see the retinal image decomposing here, and we can see a shape occurring.

Information is extracted from the visual cortex in these columns and we do several things such as stereopsis and depth perception. We can do that as the images from the two eyes now come together in this area. We have orientation lines, and as we go down, each bit of the column has an orientation in a certain way, so all 360 degrees of orientation are represented. Furthermore, there are cylindrical blobs in these areas as well as other bits of processing which go on. These are probably related to things such as colour, and there are red/green cells here that respond to local colour contrast before sending the image upstream to other visual areas, which are called the associated visual areas.

This chap worked in the lunatic asylum in the West Riding of Yorkshire. He tried to electrically stimulate the brains of animals and he then worked out that there was a map and certain things were happening in different areas. When a patient came in who was paralysed, he predicted where the tumour was, and the surgeon, Mr Macewen, went in and removed it safely. For this, he became famous. He then went on to study vision and while he did not get that absolutely right, he accidentally discovered something probably more important. He took out the occipital lobe and actually found they still had vision. As I have shown you, you have to take out every single last bit because, if you just take out the back bit, you are only taking out central vision because of the magnification. If you leave the last 5%, you are leaving 95% of the vision behind, so that is why he couldn’t show this.

He took out this area, and then found out that the monkeys could not reach for a cup of tea - they were British monkeys that drink tea, whereas, Herman Munk, in Berlin, did not offer his monkeys tea, and violently disagreed with this. Once, he was asked, “What about Mr Ferrier’s experiments?” and replied “I have not said anything about Ferrier’s work because there is nothing good to say about it.” Vicious! “Mr Ferrier has not made one correct guess. All his statements have turned out to be wrong.”

He had actually removed a visual association area, and this was the area converting vision into signals that we can then use to direct movement. We had moved beyond simply vision.

It turns out that there are two ways that images can probably be processed, and we have two streams. We have a stream down into the temporal lobe, and we have a stream up into the parietal lobe. It could be that the visual image is first processed in the visual cortex, V1, and then it passes intact through a series of filters for further processing to extract perception - that elusive goal, like awareness. We do not know where it is even though we know what it is. On the other hand, images could be broken down in the visual cortex, decomposed, and different parts of them could be sent to different individual analogue computers that specifically do one job. That is the problem with analogue computers. They are brilliant at what they do, but they only do one thing, so you need lots of them. However, that is not a problem for the brain. There are billions of cells and billions of connections in the human brain! There are more connections in the human brain than there are stars in the known universe. We do not have a problem with the number of cells. We can make as many analogue computers as we like, and that is precisely what we do. The areas are: V1, which is the visual cortex, and then V2, V3, V4, V5, and a couple of other ones which do not fall into the V.

V2 is the very next one, straight after the visual cortex and wraps around this area here. It is a map, and it is a mirror image map of the V1 map. One input comes from the visual cortex and it goes into these stripes, and then it is broken down into parallel streams. One goes up to colour, one is involved with form, and there is another stream that comes from V1 that misses it altogether and goes to this area, called MT, which is involved with motion and stereo-perception. We now have separate channels being broken down that deal with different parts of the visual image, different parts of perception. This is called parallel processing.

For example, orientation of depth, orientation of images and the depth perception is done in V3 and this comes up through two separate pathways. One is coming from V2, where a little bit of processing of something might be done, and one native, comes straight out of the visual cortex. These two different ways could be responding in these different pathways up into the parietal lobe which deals with motion in large patterns, and then the ventrolateral which deals with movement and where things are.

Analysis of colour shows that there is in another area called V4. It was discovered by Semir Zeki in London, and this receives its information, mainly, from these thin stripes, and some from the interstripes. This is involved with a different parallel stream for form. There are other areas such as V5, as we have mentioned before.

The analysis of motion is really important. It is a separate sensation. It is rather like smell or taste. Motion is not a series of still images that we add up.

The next issue is how we analyse depth. We have mentioned before that the retinal image, or any image that is a shadow, does not actually tell us about what could have made it. Any of these confirmations of these completely different structures, if the lighting is right, could cast that shadow or that image. We now know that the retinal image is inherently ambiguous, and this can be used from a number of ways, such as these tromploi. This is not really a piece of marble. They ran out of money when they built this room, so they had to paint it on. Not only that, they had to paint on the columns, which do not show well, and are actually less successful. Nevertheless, we get a real 3D perception here which is completely false. The reason we get that 3D perception is that we are making assumptions about the image and about where the light comes from.

Even though the image is flat on the back of the retina, it contains lots of information from which we can extract information about depth. It has depth perspective. The shape on the retina, as we said, is ambiguous, and perspective can suggest three dimensions if we make certain assumptions.

You can get texture gradients. Here, it shows the grass is rather coarse, and as we go up here, it gets finer and finer, until it is painted as uniform green. That is roughly how we see it, so texture gradients that are denser and coarser represent nearer things. We get shape from shading. This only works if we know where the light source is from, and for most of our evolution, the light has come from above, so these shadows are telling us where things are and about their shape. There are special cells, in the area V4, that measure the shadow below. If you rotate the image 90 degrees, they stop firing. You can get interruption of lines. Jesus’ cross interrupts the walls, so we know the walls are behind, and we have specialised cells in V1 that tell us about that. I showed you the end-stop cells earlier. Upward sloping ground tells us that these guys are further away than these guys.

Size constancy is an issue. We physiologically expand the image. These guys are drawn small and this guy is drawn big, yet we do not see these people as dwarfs – we see them as the same size people who are further away. So, psychologically, we are expanding the perception of this small image. Finally, there is atmospheric perspective. Things that are blue, due to Raleigh scattering, in the atmosphere, are further away. You can get many things out of this simple image if you look for it which is precisely what the brain does.

Shape from shading is quite important because it depends on where the angle of light is from. These are identical yet this looks like buttons and this looks like a depression because we assume the light comes from above – it is an identical image, just inverted. This is important for recognising faces. Upside down faces can be difficult to recognise, and there is something that is not right about this image.

If we do turn it the right way up, we realise how important orientation is with respect to gravity, and lighting, for our images and understanding. We have turned the mouth and the eyes upside down. We cannot tell that when we look at a face upside down, because we are so used to seeing faces the right way up – we are hard-wired to see faces the right way up. This is the Chiaroscuro light and shade technique which was used in the Renaissance to add depth and reality to their paintings. This is Gerard David’s painting on the “Rest on the Flight to Egypt”.

As we moved further into the development of the industrial age, lighting could come from below. This painting in the National Gallery actually is really rather interesting, not just because it is showing that light can come from below. Also, this is a planisphere showing how the planets and the Moon rotate. These faces represent the different phases of the Moon – here is the full Moon, and here is the Moon at either side. It is a rather interesting when you go and look at that painting and it is rather fun.

For analysis of distance you could use motion parallax, like the pigeons used to do in Trafalgar Square before the Mayor shot them all. If you ever look at a pigeon, they do this, and what they do, when they are doing this, is that this moves faster if it’s nearer, and it moves slower if it is further away, and that tells me where something is compared to this. It also tells me if it is approaching me which is why pigeons fly off if you are not carrying anything that they want to eat. We have purpose-built range detectors to judge distance.

The image changes very slightly between the two eyes, in an area here, and when it is looked at with two eyes, 93% of the population suddenly see something, in three-dimensions, jumping out at them – a square or a circle. It is actually a shift of only a few thousandths of a millimetre which changes the light on a single cone receptor by about 5% which is enough to cause it to fire. Upstream of this, individual cells, which have receptive fields in different places, can detect tiny amounts of disparity, smaller than a cone, by using this mechanism, and these parallax detectors are really important.

Motion is very difficult to depict. This is Maxim and this is Freddy, racing round this circuit, and the photograph makes them look still. More than that, this is the daguerreotype – it was the first photograph of a street scene ever, and this is crowded – it was rush-hour. The only thing you can see here is the boot man, blacking the boots of his customer who has stayed long enough for it to be impressed onto the daguerreotype. When Morse came to visit Paris, to look at this wonderful Morse of Morse Code he said, “Moving objects are not impressed.” The sense of movement is very specific and computed directly from the retinal image. If you want to show things moving, rather like this little dog running along, you have to give them have lots of legs. This is Dynamism of a dog on a lead by Giacomo Balla.

We have a movement detector. I have mentioned that it is present in the retina, and it is in the brains of monkeys and of us. What happens is that the car is racing along in this direction. As it trips this wire, it fires, and then as it trips this wire, it fires. If these images fire together, we know that it is coming along this way, and the reason we know this is that there is a delay built in called a synapse. So this fires, and is coming up here, then this one fires, and it comes up, and by the time this hits here, that one hits there, so if they fire simultaneously, we know that there is a movement, moving along this direction. We have movement detectors that also go the other way; we have movement detectors in both eyes; and the disparity ones can detect movement and tell us movement in three-dimensions of the parallax. We use that as a sort of test by delaying images in one eye, or as a natural experiment in multiple sclerosis - when the nerve becomes sick at the back of the eye, people lose this ability, this delay.

I mentioned that analogue computers are good, so there is an issue as to why we have so many maps. We have them because these things only do one thing, so we need a map of each thing we do, practically. Not all the maps are geographical maps, by any means, but they are all centred on one frame of reference, and that frame of reference is your foveal fixation. If you move off, all the maps come in to register, wherever they are in your brain, on this fixation. Many analogue computers are all working at the same time all having different functions, as in this big map room in the Doges Palace. There are lots of different types of maps, in there. Object recognition, facial recognition, colour and movement are all different jobs that are being done.

Beau Lotto and Dale Purves came up with this probabilistic model of vision. They said that we use experience to learn and we use experience for sight. They mean that we adjust the weighting of what we experience by what the likely outcome is. I mean that if something looks like a dog and it smells like a dog, it probably is a dog. It might not be, it could be a cat or a wolf, but it is very likely, in our environment, to be a dog. A binary digital computer, which is all or none, is of no use to us in this messy real world, where noise degrades the information. We need a learning computer.

This is the Perceptron, which was based on how to win at horse races. We have an outcome; it is a win. If we have no information from the weather then we do not put in any information and we wipe it off. If the form was important, and it sometimes is, we increase the input from form. The ground was not important; it was a dry day so it did not affect it. The jockey was very important so we increase the input it will have this time. The weight they are carrying around can also slow a horse down as it is going round the track. Next time, it loses, so we adjust all the things again, and if we do this hundreds and hundreds of times, we eventually end up with a very good predictor of what will make our horse win.

Now, this learning computer is using Baye’s theorem. The theorem is that we have a model called a prior. We then collect some evidence, alter the model as is needed and then develop a posterior. It is a generative model – we are generating the model according to the data. Sometimes, we get it wrong.

The Moon is about as bright as a lump of coal – it has the same reflectance. However, most people, when they look at a lump of coal, think it looks black, and most people, when they look at the Moon, think it looks white. This is because the Moon is reflecting against a background that is totally dark, and so things that are reflecting that are brighter than their background are white. Things that are reflecting lower than their background look dark, so most of the time we see coal when the background is brighter than the coal. One of the wonderful experiments NASA did was to hold up lumps of coal in a darkened room, illuminated by secret spot-lamps, and the coal glowed beautifully white, rather like the Moon.

This probabilistic strategy is based on past experience, and it probably explains this difference between what we see and reality. I have a difficulty with that Beau Lotto idea because I do not know what reality is. Reality, again, depends on a definition, but we will come back to that in a moment. Sometimes, you end up with two stories that could be true. This is the cube that we have seen before. If we look at this image for long enough, it suddenly flips, so we can either have a cube coming out at us, or we can have a cube going away from us, and it just spontaneously flips. The brain cannot make up its mind, so it chooses both. This could be a duck, or it could be a rabbit. We cannot make a decision - the two are equally valid.

Beau Lotto has a book coming out in two months’ time called “Why we see what we do” although this is not a trailer for that – I only found out the book was coming out after I put the title in last year, but it should be a very good book. He is saying that the evidence drawn from their experiments on perception of brightness, colour and geometry support the idea of Bishop Berkeley’s dilemma. His dilemma was that he did not know how we knew what was going on when the retinal image was so uncertain of what it was. He says that we make a probable guess of what it is, and that guess depends on our own experience, on what our grandparents saw, on what their grandparents saw, because if they got it wrong, and they got the shark detector wrong, they were going to get eaten. Therefore, we have evolved a very sophisticated series of complex analogue computers to analyse vision and we do it rather like the Perceptron. Every time we see something and we get it wrong, we modify the inputs, until we get it right. After we have done this thousands and countless of thousands of times, in childhood, we get rather good at it.

Now, if this is true, we should not need a retinal image to have a perception, and of course we do not. Hunter Davis’ book, “Fear and Loathing in Las Vegas”, describes these amazing images coming out of the carpet and threatening to eat him when he had slightly overdosed. They decided, in the open-top car, to take all the LSD, all the uppers, all the whisky and all the downers, in one go. They arrived at the hotel completely out of their trees, and there is this remarkable passage where he described what went on. Coleridge fell asleep after laudanum just after he had read a passage about the garden of Kublai Khan, and, of course, that led to a wonderful work of art. De Quincy observed that “If a man took opium whose talk was about oxen, he would dream about oxen.”

We met Johannes Peter Muller before. He was the son of a Koblenz shoemaker and physiologist. He wrote on fantasy images and was experimenting with hallucination. Some people, just before they fall asleep, or just when they wake, get a very vivid visual hallucination, particularly if you rather overdo the drugs, and in his time, those drugs were caffeine and alcohol. He was experimenting and sleep deprivation and caffeine led to him having a mental breakdown. He actually recovered and in April 1827, married the beautiful and talented musician Nanny Zeiller, and immediately broke down. If that is an advertisement for marriage, perhaps you need to watch the caffeine.

In summary, we have inputs from these centre surround fields. We are here, and in lateral geniculate ganglion, extracting lines. We are not doing anything else. We are extracting lines; where it is dark and where it is bright. We then process this, and then they come up into binocular cells - this means that they are fed from both eyes. These are fed from one or the other eye, they then feed into these binocular cells, and they go through separate channels, which analyse colour, form and motion, which is a separate visual sensation. We end up transforming this series of Gabor batches in V1 into a complex, beautiful, vivid, three-dimensional moving image, set in a textured background. Therefore, it is important that we do this accurately and get it right every time, so that we know what is dinner and what is a shark.

George Berkeley went quite far actually, and he said: “It is indeed, in my opinion, strangely prevailing amongst men, that houses, mountains, streams and rivers, and in fact all objects have an existence that’s natural or real, distinct from their being perceived by the understanding.” He says that nothing out there is real – it is just going on in our heads.

This wonderful limerick was then written. It says:

*“There was a young man who said*

*“God must find it exceedingly odd*

*To think that the tree*

*Should continue to be*

*When there’s no one about in the quad.”*

An anonymous response to this was:

*“Dear Sir, Your astonishment’s odd;*

*I am always about in the quad.*

*And that’s why the tree*

*Will continue to be*

*Since observed by,*

*Yours faithfully, God.”*

Of course, Samuel Johnson said: “I refute Berkeley thus”, and kicked the stone, proving that the stone was real. However, he did not refute Berkeley at all. All he did was to say that he had a sensation in the end of my toe that made him believe that there was a stone there yet it could still also have been in his head.

Ladies and gentlemen, thank you very much indeed. I hope that that has gone some way to helping us realise how complex vision is, and that it is not as straightforward as we see when we wake up in the morning and open our eyes. It is a terribly complicated process that involves over half the computing power of our brain, which is using a tremendous amount of energy, even when you are asleep and having visual hallucinations in your dreams.

©Professor William Ayliffe, Gresham College 2011