

Chapter 8

Creating the Conscious Visual Moment

A notable characteristic of the visual world is that it's a stable world. Our eyes dart about, fixating momentarily on one portion of the scene and then sliding to another, our heads turn, we walk, run and bend, so static objects are stimulus patterns that mostly move across the retinal field. Conversely, we often follow a moving object with our gaze, and the moving object remains pretty much fixed in relation to the retina. Yet what we see is a still world, with the self moving through it, and we correctly identify the other moving phenomena. This implies that some pretty dazzling calculations go into the creation of the picture.

The mechanisms responsible are found in parietal cortex. While one set of outputs from the occipital cortex goes to the inferior temporal lobe another is transmitted to the parietal lobe, to a region which is less concerned with discovering what a visual object is than with working out where it is. Here the sheets of neuronal columns don't make simple maps of the current retinal field. Eye position is brought into the calculations, and the receptive fields of the cells are determined by the direction of gaze as well as by retinal position, so they constitute a head-centred representation of the view. In later stages of the pathway input from the vestibular organs is added, allowing head movements and body movements to be taken into account. Increasingly the maps come to cover the whole potential field of view.

Damage to this parietal region can produce difficulties in perceiving spatial relationships, as in the case of an unfortunate man who suffered two strokes, affecting the area bilaterally. In one test he briefly saw an X on a computer screen and was asked at each trial whether it was to the right, left or centre of the screen, or whether it was in the upper, lower, or centre part. His performance on this simple task was only 70% correct. In another test he was asked to judge whether an X had appeared to the right or left of an O, or whether it was above or below, and he found this task extremely difficult and got only about half right, the result that would be expected from chance.

Shown the two letters in sequence and asked whether they had appeared in the same place or not his performance was equally bad when the relationship was horizontal, though slightly better when it was vertical. Incorrect judgements that the letters had occurred in the same place were made more frequently than the opposite error.

In addition to the calculations which anchor the picture in space there may be some help from the fact that, as Tom Bower pointed out, one's own nose is always in view - unconsidered, shadowy shape though it may be - and eyebrows too are usually detectable at the edge of vision. Most other mammals must equally be able to see their own noses, and those few species which have unusually flat faces generally make up for it with fringes of hair that hang down over their eyes, while flat-faced birds tend to have tufts of feathers similarly arranged. These intrusions may well provide an additional means of relating the position of the self to the view. Input relayed from the muscle-stretch receptors in the neck must also help to calibrate exactly where the head is pointing, and for some directions the sight of other bits of the body must contribute.

The parietal lobe seems to need this sort of help, for while its calculations about eye and body movements are very efficient at identifying spatial relationships, they don't by themselves create a wholly stable world. If someone sits in a completely dark room observing a stationary light, the light soon begins to appear to move. Moreover if, after a short while, the observer is asked to point to the light, their gesture is often inaccurate. Equally odd things happen if a subject in a dark room is shown a luminous dot in a luminous frame. When the dot moves and the frame is stationary the movement is correctly identified. But when the dot is stationary and the frame moves it's usually still the dot that is seen as moving. On the other hand, if instead of the dot there is a small luminous shape that looks like a ship or an aeroplane the observer may see it moving forward, but is less likely to be deluded into seeing it travelling backwards.

It seems that we don't depend exclusively on knowledge of eye movements and body movements to know what's static and what isn't. We also make use of the expectations derived from previous experience, and treat certain things as likely to be fixed landmarks - anchors for the 'picture' - and others as likely to move. Information from the inferotemporal pathway is used to augment the calculations in the parietal pathway. Deprived of fixed landmarks we soon become confused, which is why it's so easy to get lost in a featureless desert.

When the temporal or ventral and the parietal or dorsal pathways were first distinguished the former was described as a 'what', the latter as a 'where' pathway. However the operations carried out in the parietal lobe have turned out to extend well beyond locating the stimulus patterns that are being identified in inferotemporal cortex. The further calculations needed for preparing actions on them are also performed there. One subdivision is involved in managing eye movements, another directs the reach towards the object, a third calculates how the hand must be shaped to grasp it, and a specialist sector of the latter organises the appropriate grasp for a tool that is to be used rather than just picked up.

Just what the visual areas of parietal cortex can do has been demonstrated by a study of a patient with a lesion in inferotemporal cortex that left her unable to recognise visual objects. Nevertheless she shaped her hand correctly when she grasped something, and when she was given a rectangular object to pass through a slot she orientated it in the correct way to go through the opening. Clearly the parietal cortex receives quite a lot of information about shape. Indeed, the parietal cortex is more objective about visual objects than the temporal. Our conscious experience is sometimes subject to illusions – here, for example, the circle surrounded by smaller circles looks larger than the one surrounded by larger circles, although they are actually both the same size. Even when the conscious experience is deluded by such arrangements, however, the hand put out to pick up the object gets shaped to the appropriate size.

This evidence suggests that the ventral visual pathway's function might best be described as deciding whether a stimulus pattern should inspire any action, and the dorsal's as working out how the action should be performed. Establishing just where the stimulus is constitutes only one step in the latter process. However, from the point of view of creating the picture in the brain it's combining the 'what' and the 'where' that's important. The necessary connections between the two pathways are gradually being uncovered.

The unstable eye

Saccades and tracking movements are not the only motions of the eye. The fixations are not held steady; the eye regularly drifts a little, and then flicks sharply in a microsaccade - enough to shift the image over several receptors. This movement can be reduced with practice, but not wholly eliminated. You can observe the effect when you've looked at a bright light for too long and experience an after-image. Being fixed on the retina, the after-image trembles and dances along with it. The picture-in-the-brain shows something that is absolutely fixed in relation to the retina as moving.

One obvious means of exploring the significance of the mini-movements is to see what happens to perception when their effect is eliminated. Back in the nineteen fifties an ingenious device was developed to achieve this, a contact lens with a little stalk attached, to the end of which a tiny, simple line drawing could be fixed. It turned out that a stimulus thus stabilised in relation to the retina might quite rapidly vanish. Within a period of as little as two or three seconds, it was reported, the lines of the drawing would soften, reducing the contrast between line and background. Almost immediately some part of the figure would disappear, and sometimes the whole figure, the space being filled in by 'background'. The disappearance was not permanent, however. The lost bit of drawing would re-materialise within seconds, while another bit would vanish in turn.

By the nineteen nineties images could be stabilised on the retina more comfortably, with computer displays that were constantly adjusted for microsaccades, and it was confirmed that lines do tend to disappear in these circumstances. The thinner the lines are and the lower the illumination

the speedier the vanishing act. Images defined purely by colour, with no luminance contrast, behave in the same way. A stabilised green disc against an orange background disappears very quickly. And something particularly intriguing happens when - as Crane and Piantanida ingeniously accomplished - the boundary between two coloured stripes is stabilised on the retina, while the outer borders are unstabilised against a dark field. The boundary vanishes within seconds, and the experimental subjects reported seeing dots of both colours, or islands of one colour against a background of the other, and in some cases a mysterious, indescribable combination of the two colours. Finally, in 2007, Michele Rucci and colleagues, using a very sophisticated technique of image stabilisation, confirmed that fixational eye movements do help with the perception of fine lines (high frequency gratings), but not with thicker ones.

Collewijn, Martin and Steinman, meanwhile, tried a different trick. They experimented with set-ups that gave their subjects the chance to compensate for an abnormal amount of movement of the retinal image, in a way that would have made it possible to achieve perfect stability. They found that the degree of compensation chosen did not maximise stability, but maintained a certain appreciable degree of image movement.

So the eye's little drifting movements do perform a useful function. Fine detail is registered more effectively with their assistance, resulting in a better picture-in-the-brain. Further experiments by Rucci's group have shown that this is particularly useful when viewing the complex small-scale detail of natural scenes.

Part of the explanation may be that minisaccades serve to move a feature across the receptive fields of several retinal ganglion cells, several lateral geniculate cells and several of V1's feature-sensitive neurons. This increases the chances that at some point it will be perfectly centred on the receptive field of the appropriate sort of neuron and cause the maximum rate of firing. As a result population coding must work more powerfully at each level, making use of the reports from a larger number of cells than would otherwise be the case. The amount of information that can be obtained from a limited population of feature-sensitive neurons is thus increased. The firing rates of a whole clutch of V1 neurons are in effect compared over the duration of the not-quite-steady fixation, and a more reliable conclusion about the feature is reached than could otherwise be achieved.

This system should help to compensate, too, for the imperfect optics of the eye. It might be regarded as not unlike the astronomers' trick of superimposing several pictures of the night sky, one over the other, in the hope that the light which shows on all of them indicates a genuine star, while the fuzz caused by reflections off particles in the earth's atmosphere will vary from one exposure to the next and can be ignored. The drifts and microsaccades of the eye allow for dirty data to be averaged, and thereby cleaned up. Such a system might well be misled by an edge or a fine line that was stabilised in relation to the retina, and thus excited fewer feature-sensitive cells than normally happens and produced a lower overall rate of firing. From time to time such an input might be judged to be 'noise' rather than information.

In addition the constant shifting of the image on the retina means that cells which are excited only by moving stimuli, both retinal ganglion cells and complex cells in primary visual cortex, can contribute to the business of registering stationary ones. Several different types of response have been found in the cortical cells. One sort fires when a microsaccade causes a stimulus to move into, or across, or out of the receptive field, but only when the movement is in a certain direction. Another behaves similarly, except that it isn't direction-sensitive. Others, with rather smaller receptive fields, are activated only by stimuli that remained centred. The fourth sort shows a burst of activity when a stimulus is brought into the receptive field, and continues to fire at a lower rate during the period of stabilisation and slow drift. The third and fourth of these categories clearly respond optimally to stationary stimuli. The other two, especially the neurons which are sensitive to movement in any direction, must provide a means of extracting information about non-moving objects as well as moving ones, thanks to the eye's micro-movements.

From a subjective point of view it's easy to see a static stimulus pattern, more difficult to see one that's moving, and perhaps confusing the experience by changing its shape as it does so, or going too fast to be seen as anything more than movement. Stationary objects tend to provide a much more satisfactory visual experience. However, from the perspective of creating a neuronal mechanism to deal with its movement must be the easier proposition. A moving object obligingly defines itself by activating a whole row of retinal ganglion cells, of lateral geniculate cells, and of V1 feature-sensitive cells. Its passage thus provides a good deal of data at all levels of the machinery, and if some of the cells fail to operate it doesn't matter because the movement of any object of more than minuscule size will have activated a good many of them, more than enough to trigger the direction-sensitive cells of V1 and MT. (Movement-detectors have to operate pretty efficiently, of course, since they only get one crack at the stimulus.)

A stationary stimulus is another matter. Excitation in a bundle of feature-sensitive cells could indicate a stationary stimulus, but could equally be the result of a moving stimulus passing through their receptive fields. In order to discover which it is, it's necessary to establish that the stimulus being registered now is the same one that was being registered a moment ago. For some species the solution is no doubt to wait a moment and act only if an adequate level of stimulation builds up over that period. The system which compares a series of inputs as the eye drifts and then flicks back towards its original position is subtler. It exploits the necessary delay to extract finer information from the data collected. The picture-in-the-brain is thus the product of what might be called a comparing memory, whereas the animal that merely needs to accumulate a certain amount of input before making a response to a stationary stimulus uses an accumulating memory.

A parallel can be drawn between vision and the tactile sense. We learn a great deal more about an object by running our fingers over it than by just touching it. Rats and mice use a similar tactic when they are exploring - they twitch their whiskers forwards and backwards, at a rate of around eight cycles per second (a gesture known as whisking). Clearly more can be learnt from a changing input, achieved through a displacement of the receptors, than from a static one. Vision works in much the same way, except that the eye is continuously in exploratory mode, and makes use of very tiny movements as well as larger ones.

For us the benefits of microsaccades are only modest, helping us to see fine lines in poor light, or to sort out potentially ambiguous clusters of features. For species with fewer photoreceptors, fewer ganglion cells and a smaller primary visual cortex they should be more conspicuously beneficial, especially for species which undertake most of their activity in dusk or darkness. I suspect that microsaccades explain why mammalian eyes look so different from those of reptiles, amphibians and fish. For species with foveas they must be particularly useful for getting significant lines and edges centred precisely on the small receptive fields connected to that bit of retina.

Creativity takes time

Another significant aspect of the picture-in-the-brain is that it takes time to create. We seem to perceive the world instantaneously, but a measurable amount of time is needed to collect the data and turn it into conscious perception. Once created, though, the conscious experience has a life of its own, and may outlast the input. The relatively clumsy pace at which conscious vision operates accounts for the fact that a movie running at 28 frames per second is perceived as showing smooth, continuous action - though the early 24-frames-per-second looks a little jerky.

How much input is needed to produce a conscious experience, how long does it take to create the picture, and how long does it last, once created? The answers have been very revealing.

First of all, the input. The tiniest, briefest flash of light - just a few photons delivered to a dark-adapted eye - can produce a conscious experience. But experiments with the tachistoscope show that for any sort of pattern or shape to be experienced a minimum of ten milliseconds of well-illuminated input is required. And that's when the stimulus pattern is a very simple one, such as a letter, or a well-defined line drawing consisting of only two or three clearly related features, and where the contrast is good so that averaging the inputs over a longer fraction of a second isn't necessary.

Considerably more time is needed for a complex pattern to be turned into a meaningful conscious experience, as some of the experiments described in the last chapter demonstrated. The time required to collect data explains why a fast-moving object is seen only as a blur - it passes across the retinal field too fast to provide an adequate sample of stimulation, and as far as the conscious experience is concerned merely interrupts the collection of evidence about the scenery behind. If the object is travelling fast enough (a bullet, for instance) it doesn't even do that.

After an adequate tachistoscopic exposure the 'picture' doesn't appear immediately, but takes a further fifty or sixty milliseconds to materialise. The fairly precise measurement is inferred from the fact that if a second stimulus pattern is registered on the retina within this time the first one doesn't reach consciousness. The deduction is that the sensory input has to be processed to turn it into conscious experience, and an interruption by further input disrupts the operation.

The conclusions reached from experiments with the tachistoscope fit well with the timing of the eye's saccades from one fixation to the next. There are generally between two and four fixations per second. The saccades in between take about 25 milliseconds each, on average, so fixations must generally last from about 200 to 500 milliseconds. On the evidence of the tachistoscope, this is about the amount of time needed to collect data, process it, and appreciate the result. Complicated scenes require fixation durations at the longer end of the range, but for dealing with simple, clearly defined and well-learned stimulus patterns such as printed words 200 milliseconds can be adequate.

What also happens during the fixation is that a brief record is established. We know this because if the scene viewed is succeeded by a blank visual field the 'picture' can linger on after the retinal stimulation that produced it has ended; on a dark visual field it can last for a matter of seconds. You may have experienced the effect under natural circumstances, during a thunderstorm on a dark night. A flash of lightning illuminates the scene that was invisible before, and the visual sensation fades slowly, lasting longer than the illumination provided by the lightning. In more usual circumstances the record must serve to promote the sensation of a static scene when the eye is saccading and what's being registered on the retina is a blur of movement. The illusion is helped by an arrangement which during a saccade automatically suppresses transmission in most of the channels leading to visual cortex.

The mechanism that creates the lingering image was named, by Ulric Neisser, iconic memory, while the image itself came to be termed the icon. The seminal experiment on the icon was done by George Sperling. He showed his subjects a tachistoscopic exposure of a three-by-three array of letters, and asked them to read out one line, but they were only instructed which line to read (by means of a high, medium or low tone) after the exposure was terminated. As long as the signal indicating which line was required came immediately after the flashed exposure of the letters they read the letters off quite easily, and reported that they were still experiencing a conscious percept when they did it.

Other workers developing the possibilities of this experiment used only a single line of letters and asked their subjects to read out just one, and this time the required letter was marked by an indicator flashed on the screen immediately after the display of letters. This turned out to work as expected as long as the indicator was just a small line underneath the space where the chosen letter had been. In the subjective visual experience this translated into an actual underlining of the letter. However, when the experimenters attempted to mark the letter with a circle enclosing the space it had occupied things turned out rather differently. The subjects saw the circle, but not the letter. The sign which was meant to indicate which letter should be read out actually made it invisible. This is how it was discovered that new input can block out preceding input if it follows too fast.

An obvious question to be asked about the icon was how it relates to the retinal image and to the big picture. Is it fixed in retinal space or objective space? Since conscious vision gives us an objective view of space the latter might be expected. A variation on that last experiment was devised to elucidate the matter. Once again a row of letters was flashed on the tachistoscope screen, and once again a circle followed, positioned so that (intentionally this time) it would obliterate one of the letters. The new wrinkle was that the subjects were trained to centre their gaze on one fixation point,

and then to transfer it very precisely to another, nearby fixation point before the circle appeared. It was theorised that if the position of the icon was fixed in relation to the retina then the circle would replace the letter that now occupied the same bit of retinal field. If the creation of the icon took eye movements into account then a different letter would be masked.

As it turned out, neither prediction was fulfilled. The letter which was missing was the one that occupied the same site on the retina as the circle. But the circle did not appear surrounding that empty space. Instead it was seen framing the letter which it was expected to mask according to the icon-related-to-alloentric-space hypothesis. It must be concluded that the calculations that turn retinal input into a picture that depicts objective space are efficiently done before the conscious experience materialises, even under very strange conditions.

Other workers confirmed the finding a few years later with a slightly different experiment. The subjects saw a tachistoscopic exposure of a pattern of twelve dots. They then shifted their gaze slightly to centre on a new fixation point and a second array of twelve dots was flashed on the screen. The two collections of dots were designed so that, overlapped in one way, they would make up a regular array with one dot missing, and it would be easy to identify the gap. Superimposed in the other way they would be just a patternless jumble. The result was the same as before. The two displays were combined in the way that produced the correct relationship in terms of allocentric, objective space.

Clearly one of the operations carried out in the period before the visual scene reaches consciousness combines the results from inferotemporal and parietal pathways so as to locate what is seen correctly. Other processes are sometimes completed successfully in this brief period of time, sometimes not. If the stimulus pattern is a familiar one it evokes excitation in a clear-cut set of finely tuned neurons in inferotemporal cortex, as described in the chapter on learning to see - a process which is assumed to equate with matching a visual input to a visual schema. When that happens a clear conscious experience can emerge at maximum speed. A novel stimulus pattern, for which new neurons have to be trained, is likely to take a little longer, so that an initial fuzziness in the experience is detectable - if one stops to think about it. Wholly novel and strange objects viewed under difficult circumstances, may, as shown in the previous chapter, produce a very muddled impression until they have been thoroughly studied, and a satisfactory interpretation of the data discovered.

Another process that may be successfully completed during the initial stage of creation but may take longer is that which establishes the third dimension of the picture. We saw in Chapter 6 how the binocular disparity cells of primary visual cortex function to establish the distance of visual objects. However, there is an obvious hitch to overcome before they can do their job. For the binocular disparity cells to achieve their function the convergence of the eyes must be correctly adjusted; but before the eyes can be correctly converged, not to mention focussed, it's necessary to know how far away the object under inspection is. If the focal plane needs to be altered for the new input the problem can only be resolved by a dialogue between the focussing and convergence mechanisms and the binocular disparity detectors - in other words by a little experimentation. In most cases the binocular disparities recorded in the previous view will provide the basis for a fairly good guess about how much convergence and focus have to be altered to shift the gaze to a nearer or more distant object. But the result is only likely to be exact in familiar situations. On other occasions some further juggling may be necessary before the best possible match of the two views is achieved and the object of interest successfully engages the zero-disparity cells.

We can deduce that this dialogue generally takes place before the conscious visual experience emerges because it can be detected when, occasionally, it continues for longer. This happens when the scenery is not perfectly obliging in arranging itself so that there are no ambiguities. Where groups of identical features are clustered closely together pairing them up in the two retinal images must be tricky; if the convergence of the eyes is marginally wrong there are likely to be quite a few features which can be inappropriately paired in the receptive fields of binocular disparity detectors. On such occasions it must be necessary to sample the images through more than the usual number of convergence adjustments. Perhaps the eye's tiny drifts and microsaccades help with this process,

perhaps they make it more difficult. But certainly one may stare at a complicated bit of scenery for some time before working out the exact ins and outs of it. The effort involved becomes particularly detectable when one looks at a random dot stereogram in a stereoscope. At first sight it usually appears to be a simple two-dimensional array. It's only after sustained study and perhaps conscious effort that the three-dimensional figure suddenly pops out - but once it's discovered it's easy to maintain the perception.

How the conscious experience is created

An increasing amount of evidence supports the long-held idea that the conscious sensory experience equates with an extended circulation of impulses around a complex collection of circuits. Distant brain areas become interconnected temporarily, to work together. A certain minimal amount of input is needed to activate all the circuitry, and a certain amount of time – just how much depending on the complexity of the input, and the degree of familiarity.

In the visual system some of the circuitry must serve to link inferotemporal 'what' information to parietal 'where' information. Other connections provide for visual input to be linked to any coincident auditory and tactile input. Visual, auditory and tactile pathways all continue to the frontal lobe, where the information about the outside world can be combined with information about the inner world and the really important conclusions about sensory input are reached, namely what action it should elicit.

The various stages of the cortical sensory pathways don't just pass information forwards, however, they also send messages back to every station. Some project back to the lateral geniculate nucleus of the thalamus, where fibres delivering information from visual cortex far outnumber those carrying it in the other direction. Some connections even run back to the retina. Thus all the later stages of the pathways can influence what goes on in the earlier parts. The creation of the conscious experience can be imagined as involving a dialogue between numerous cortical and subcortical areas. The inferotemporal neurons activated by the current batch of input send messages back to the occipital lobe which serve to intensify the activity of selected neurons in earlier stages of the pathway, so as to fit with the likely interpretation of the retinal input. In the thalamus neurons fire more strongly when messages returned from visual cortex indicate that they are recording an object, rather than background.

For the most part this process makes for rapid and efficient identification of significant objects. Occasionally it leads to mistakes, making us see what we expect or hope for rather than what is actually there. And clearly it's what makes it more difficult to try out an alternative interpretation of the data once an incorrect guess has been made.

In addition to the guidance from the schema-areas there is also a strong influence from attention. For instance V4, heavily concerned with sorting out colours, is strongly activated when colour is relevant to the current task, less so when it's irrelevant. When it's the direction of movement that must be noted activity intensifies in the temporal area specialising in visual movement, MT. Even in V1 the level of firing is affected by whether what's being reported in the neuron's receptive field is of significance to the business of the moment.

Some 'top-down' communications, meanwhile, serve to clarify the data. One study investigated how the occipital responses to bars moving across a chequered background were affected when MT was inactivated. It turned out that, without feedback from MT, many neurons in V2, V3 and some layers of V1 showed reduced responses. The reductions were particularly marked when the contrast between bar and background was small, so this feedback must play a particularly important role when only limited information is available.

The same circuitry also seems to help in emphasising the ends of lines. Many lateral geniculate neurons are more responsive to the end of a bar crossing the field than to a continuous one. Feedback from MT increases this distinction. The consequent intensification of the feedforward messages presumably makes line-ends and corners more salient to primary visual cortex, and helps the ventral visual pathway with the identification of the stimulus pattern. Feedback from areas such as MT,

Adam Sillito and colleagues suggested, *can inform the earlier levels about the broad borders of an object in ways that are [otherwise] inaccessible to the circuits at these levels.*

Other inputs, some facilitatory, some suppressive, come from neighbouring neurons, so that the response to a stimulus within a cell's receptive field is modulated by what's going on just outside it. McManus, Li and Gilbert suggested that the horizontal connections in primary visual cortex provide a range of possible interactions, and that connections from more forward cortical areas determine which are put to use. This helps in sorting out figure from background. As a result of all the accumulating contributions the V1 responses to a stationary, unchanging stimulus pattern evolve over a few hundred milliseconds.

The sustained circulation of impulses around a large circuit must incidentally help to iron out any problems arising from the different rates of transmission in the three separate pathways from the retina. Messages from the parasol/magnocellular system, travelling in thick axons, must arrive first, bringing information about more widely spaced features, which is in line with the fact that the overall layout of a scene is perceived first. The input from the midget ganglion/parvocellular system is needed for the details. Messages from the koniocellular blue/yellow system travel slowest, though there is perhaps some compensation in the smaller amount of processing they have to undergo on the way. But it seems possible they add subtlety to the colouring only after edges and other contrasts have been established.

Meanwhile the conversations between the various stages of the visual pathway form only part of the story. There are also multiple connections to lower areas of the brain, some of which are essential for any sort of consciousness. It's the co-ordinated flow of impulses around all these routes that creates the effect we know as sensory consciousness, a flow that can be briefly self-sustaining.

The current evolves as it goes, winnowing more information out of the original data in the process (apart from the odd occasion when misinformation results). In the chapter on the basic mechanisms of vision a set of neural arrangements was described that's often referred to as a set of filters, each extracting one particular sort of data out of the input. As knowledge of the visual system has improved it's become clear that it's a very dynamic set of filters, with feedback producing rapid modifications at every level.

By the time we are adult a great deal of what we see can be identified pretty rapidly, at least to the degree of accuracy necessary for our needs or purposes. If a stimulus pattern is a familiar one, in a likely context, the classification will be made very promptly, probably at the first fixation and before the conscious experience occurs, so the perception will immediately be clear. That recognition of familiar stimuli can be that fast is demonstrated by the tachistoscopic experiments. It's also implied by the fact that the same pattern of saccades tends to be repeated whenever an object is re-encountered, since the object must have been identified for the appropriate scanpath to be activated. The fact that we tend to be jolted into attention when the familiar scenery of our lives manifests some alteration, even though we have been scanning it for some quite different purpose, is also suggestive. The matching-to-schema process can be very efficient.

Indeed, it can be done without consciousness. Cues flashed on a screen too briefly for conscious perception can influence the subsequent interpretation of ambiguous words or scenes, or the judgement of a scene or a face. Such subliminal cues can also evoke emotions, and produce a galvanic skin response. It's only with novel or indistinct or unexpected input that the classification process takes more time than the process that leads to consciousness.

However, even the most minimal conscious sensory experience only occurs if the circulating messages reach the frontal lobe of the cortex, as Raphael Gaillard and his colleagues showed, with the help of patients who had had electrodes implanted in their brains in preparation for surgery. On a computer screen the subjects were shown first a 'mask' consisting of a row of hatch marks, then a very brief exposure of a word, then either a blank screen or another meaningless mask, the whole sequence taking only half a second. There was no awareness of the word when a mask followed it, although activity was seen in several cortical areas, including the visual cortex and parts of parietal

cortex. This died out quickly without propagating further. Without the mask the word couldn't be read, but the subjects were aware of seeing it, and could even indicate whether it was a threatening word or not. The conscious sense of something there, sometimes accompanied by an emotional response, came with sustained activity across a larger number of cortical areas, and with a notable activation in the frontal lobe.

More recent work by Pietro Avanzini and colleagues suggests that for a fullscale conscious experience the reactivation of visual cortex by returning messages is also necessary - in other words a complete activation of the circuit. Antoine del Cul and colleagues, meanwhile, have confirmed that a fair amount of processing is done in the occipital and temporal lobes before the conscious experience emerges, and that conscious perception correlates with more extensive activity that also takes in parietal and frontal areas.

An early insight into the development of the conscious sensory experience emerged from a tachistoscopic experiment performed by Haber and Hershenson. After a very brief exposure of a seven-letter word their subjects saw nothing. When the same image was flashed again, after not less than 8 seconds, they would begin to see parts of letters, or perhaps a whole letter, on the second or third flash. Perception of the whole word might require several more flashes. But when the percept developed it was in no way hazy, and the subjects were always quite certain of it, though unable to see anything a few exposures earlier. Given what's known now we can deduce that the succession of inputs builds up the efficiency of transmission in the activated pathways so that the message eventually reaches frontal cortex.

The early findings were linked to the modern ones in an experiment by Moshe Bar. Pictures of various familiar objects were presented very briefly (for 26 milliseconds), sandwiched between two masking events, with each object appearing up to six times in the course of the presentation. The subjects were asked to press one of four buttons to indicate how well they could identify the object, while an fMRI scan was performed. Successful recognition of an object was achieved after several exposures, as activity in inferotemporal cortex gradually increased and propagated forward.

Evidence from touch

The first evidence that sensory consciousness involved expanding the excitations that resulted from sensory input into a longer activation of extensive neuronal circuits came from Benjamin Libet, and involved a different sensory channel. Back in the nineteen seventies he recorded activity in the somatosensory cortex of conscious humans, working with patients who were to receive surgery for epilepsy. (It's customary, before operating on the brain, to explore the relevant area by means of gentle electrical stimulation.) Libet found that a single weak impulse of electrical stimulation applied to the skin of the hand was sufficient to cause a tactile sensation. However, when the current was applied to the somatosensory cortex it took a half-second train of impulses, in a range between 20 and 120 per second, to produce a similar conscious sensation. Prolonging the stimulation beyond the half-second mark created the impression of a prolonged tactile experience, rather than increasing the perceived strength.

Libet also reported that there was a delay of up to half a second before the subject consciously felt a stimulus on the hand. It's been calculated that only a small fraction of a second is needed for the transmission of a signal from hand to cortex, so the rest of the time must be used for cortical processing and elaboration of the message. A further interesting discovery was that a stimulus to the skin could be either masked or enhanced by stimulation applied directly to the cortex within the same fraction of a second - the difference depending on the precise temporal relationship between the cutaneous and the cortical inputs. All this showed not only that it takes time to create a conscious experience, but also that the messages arriving from sensory receptors are considerably augmented and prolonged in the neocortex.

The time taken to set up the circuit explains why a visual input is blotted out by another that follows very fast; the second input interrupts the circulation of impulses initiated by the first. It also helps to explain another phenomenon. When someone is paying attention to one sensory event they tend to miss another one that follows very closely after. Experimental subjects asked to press a button every time they see a certain word in a rapidly flashed list usually miss an example that occurs too soon after the previous one. Attention to the first word, and to the action, can be deduced to maintain the first active circuit long enough to prevent a second one being initiated.

Taking time to create the conscious sensation is fine for static scenery, but surely it's a little dodgy when it comes to moving stimuli, which may require rapid responses? Luckily the conscious experience of movement materialises a little faster than that of stationary stimulus patterns. This is shown by an illusion. When observers watch a moving stimulus and a static stimulus is flashed briefly at the precise moment the moving one is level with it, the static stimulus is seen as appearing just after the moving one has passed. Just how far after seems to depend on the brightness and salience of the static figure.

Perhaps we shouldn't be surprised. The parasol/magnocellular pathway, geared for detecting movement, uses very fast-conducting fibres. Furthermore, collecting enough information to demonstrate the existence of a moving stimulus is relatively simple, and there are numerous neurons involved at every level to ensure it's done efficiently. Establishing the existence of a stationary stimulus pattern, on the other hand, may be supposed to take a little longer.

If our conscious perception of movement is nevertheless slightly delayed there are several factors which must combine to reduce the disadvantages. MT (the temporal area which specialises in visual motion), receives direct connections both from the lateral geniculate nucleus and from another subcortical structure that receives direct retinal input, the superior colliculus, where there are neurons particularly sensitive to moving stimuli. MT in turn communicates directly to the parietal pathways which calculate the 'where' of a stimulus pattern. So it's probable that the motor system is already being primed for use before the conscious experience has materialised. These arrangements must account for the way one sometimes finds oneself tensed and preparing to duck or to throw up a defensive arm almost before one becomes aware of the need for action.

Another helpful gimmick is that we can extrapolate the future path of a moving object, which enables the response to be calculated in relation to where the object is about to be, rather than where it is now. The calculations, based on what is consciously seen and on past experience of moving objects, take into account the delay in conscious sighting of which we are unaware.

Blindsight

One part of the cortical visual pathway that's pretty much essential to conscious vision is V1. People who suffer damage in later stages of the pathways have specific problems. They don't see in colour, or they don't perceive motion, they have difficulty knowing where objects are, or they fail to recognise particular sorts of stimulus pattern - but they do have a 'picture' even if it's startlingly deficient in some respect. People with any significant degree of damage to V1 generally no longer have a conscious experience in the part of the visual field covered by the damage. They may still have some sort of access to visual information, however, for if they are asked to guess where a light or a moving object is they may perform at well above chance levels. They can often avoid obstacles too. Some of them can also guess surprisingly well at which of two orientations is being presented, or which of two colours. The ability is generally termed blindsight, or sometimes Riddoch's syndrome, after the neurosurgeon who first studied it.

Part of the importance of V1 must lie in the fact that it provides the major route by which the output of the lateral geniculate nucleus is forwarded to later parts of the visual pathways. But there are some direct connections from the lateral geniculate nucleus to visual areas beyond V1. Movement seems to be the stimulus most likely to be detected, so the short, subcortical route to MT probably makes an important contribution to blindsight. There are also connections from lateral

geniculate to V2, V3, and V4, which have been shown, in the macaque monkey, to be adequate to produce some activity in those areas.

Blindsight patients usually report that there is no conscious visual experience, they just have a sort of feeling about the stimulus. Some, however, have said that on some tests they have been aware of something - a something often described as a shadow, perhaps as a dark shadow on a dark ground. One subject who has been studied at length has said that he does sometimes have vague conscious sensations of brightness, or of visual movement. Presumably the latter is something like the sense of movement one sometimes gets at the edge of vision, without having any idea of what it was that moved. Perhaps we might deduce that a modest amount of retinal data reaching the dorsal 'where' pathway, which needs to be very sensitive, tends to have a greater impact than a similar input to the inferotemporal 'what' pathway. Sometimes, however, enough information may reach the latter to provide a hint of the stimulus.

There seems to be little consistency about these conscious sensations - the same stimulus, under the same conditions, can lead to awareness on one occasion, but none a few minutes later, for no detectable reason. Presumably it's a matter of chance whether each cortical neuron in the surviving parts of the pathways fires in response to the modest input, and consequently whether the result tips the scale sufficiently to trigger the circulation of impulses that creates some consciousness.

Blindsight is more likely to occur in people who have been cortically blind for some time, which suggested that the connections from the thalamus and other subcortical areas to parts of the cortical visual pathway beyond V1 might have become stronger than is normal. This has been shown to be the case both in monkeys and in someone whose V1 was destroyed on one side by an accident at an early age.

Recently it has proved possible to create a temporary condition of blindsight in sighted subjects by applying a brief pulse of transcranial magnetic stimulation to V1. If the pulse is presented at just the right moment it blocks out conscious perception. But these subjects too could often identify the orientation of a bar or the colour of a patch. Like the blindsight patients the reported that they were aware of having seen something, although they had no conscious experience of colour or orientation, and were only guessing at it. Transcranial magnetic stimulation to V2 can similarly prevent conscious visual sensation.

Investigations of what normally happens in the visual cortex of seeing subjects bear out the implications of blindsight. Ress and Heeger used fMRI to study activity in V1, V2 and V3 while their subjects tackled the task of detecting a brief appearance by an indistinct stimulus. They found that activity in these areas was higher when the subject reported seeing something, regardless of whether there actually was something to see. False alarms, in other words, correlated with higher activity than missed stimuli.

Possibly the early stages of visual cortex constitute not only the main means of access to the later stages but also a crucial node in the visual circuits - the place through which all the circulating messages must run if a flow adequate to create a full conscious visual experience is to be achieved. V1 and V2 are the areas with the smallest receptive fields, the ones that record all the details of the visual input, and their exact location. The receptive fields of the 'where' pathway's cells are larger than those of V1, those in the later stages of the 'what' pathway even larger. Moreover the latter pathway is concerned, the evidence suggests, with registering the essential aspects of stimulus patterns, so that they can be recognised sufficiently well for a decision as to whether they require action. Inferotemporal neurons don't seem to be very fussy about detail, except where they need to be to make useful discriminations. Indeed, there can hardly be room to record all the details of all the infinite variety of objects that may at some time or other be encountered.

V1 and V2 may therefore be the only areas that register the full detail of the scene the retina is currently reporting. As Jean Bullier has put it, maybe *the lowest cortical stages act as 'active blackboards' for integrating the results of computations carried out in the different cortical areas of the visual system.*

The binding problem

As the workings of the sensory systems have been elucidated some neurologists and philosophers have puzzled over the question of how, if the neocortex processes sensory input through all those separate channels, do the separate bits of information get put back together to create a meaningful whole? This has become known as 'the binding problem'. It may be illusory. The maintained circulation of impulses around extensive circuitry means that all the threads of evidence, even if they don't reach the final stations in frontal cortex with precise synchrony, will be brought together in a temporal overlap. This must allow the data which has been analysed in different ways through parallel channels to be reassembled into a meaningful whole. In short, sensory consciousness is an ultra-short-term memory which provides the answer to the binding problem.

This conclusion is supported by an experiment carried out by Anne Treisman and Garry Gelade, which showed that while a conscious sensation can be created out of a very brief input the minimum needed for a successful binding operation may be slightly larger. The subjects were shown very brief tachistoscopic displays of coloured letters and digits, and their attention was directed to either the letter or the digit by asking them to report its identity. Questioned subsequently about the colour of the letter or digit the subjects sometimes got it wrong. After a presentation of a red letter and a blue digit they might confidently report the opposite.

Also relevant, perhaps, to this question is the experience of the unfortunate man who suffered major damage in both parietal lobes, causing difficulties in locating objects. He also had difficulty linking form and colour. Presented with a display of two coloured letters, and asked to report the name and colour of the first to appear, he reported incorrect combinations in 13% of the trials, even when the displays lasted as long as ten seconds and his attention was wholly concentrated on the task. This indicates just how important the contribution of the parietal lobe must be in knitting the various strands of evidence back together again to build the picture-in-the-brain. The separately processed aspects of the image must be reunited largely by reference to the position they occupied on the retina. The union must be achieved by linking the parietal 'where' pathway and the temporal 'what' pathway into one circuit, and this is something else that takes a little time.

These experiments confirm that the machinery of sensory consciousness not only allows information from an assortment of different sensory channels to be correlated; it also provides the means of analysing the input of a single sense through multiple channels simultaneously, and then assembling the results. In short, the conscious moment is the manifestation of a process which brings a great deal of data together, and harvests a great deal of information from it. By means of the mechanisms of sensory consciousness the moment's information is then illuminated further by relating it to succeeding inputs, gathered over seconds or even minutes. The conscious moment is subsumed into a stream of consciousness. This process will be the subject of the next chapter.

Recapitulation

It takes a measurable amount of time to turn input at sensory receptors into a conscious sensory experience.

In the case of vision the delay serves to ensure that stationary objects are firmly distinguished from moving ones.

In addition several profitable operations are carried out during the delay.

A succession of excitations in feature-sensitive neurons is compared as the eye drifts and microsaccades, which makes for more accurate identification of poorly defined lines and edges, and allows retinal ganglion cells which specialise in detecting movement to contribute to the discovery of stationary ones.

A dialogue between occipital cortex and the visual schema areas of inferotemporal cortex leads to the identification of familiar objects or scenes.

A very brief exposure of a familiar stimulus pattern can trigger an unconscious identification which influences the interpretation of subsequent inputs.

It takes more input to make a conscious visual experience.

Still longer study is needed to make sense of unfamiliar or ambiguous stimulus patterns.

Meanwhile the visual information has been transmitted not only to inferotemporal cortex but also to parietal cortex.

There a succession of processing stations has taken saccades, head movements and whole-body movements into account, and translated the retinal image on to a body-centred map.

Thanks to these calculations our conscious experience can show stimuli that have been registered across a succession of retinal receptive fields as stationary.

A conscious experience only emerges if the excitations are strong enough to reach frontal cortex and initiate return messages from there.

The importance of the sustained circulation of impulses which creates sensory consciousness is that the extended analysis of sensory inputs can be very delicate, and a great deal of information can be brought together.

