

Chapter

NEURONAL REFLECTIONS

Rafael Malach*

Neurobiology Department,
Weizmann Institute of Science Rehovot, Israel

ABSTRACT

The search for the link between brain function and conscious awareness poses a profound challenge to Neuroscience research. While a complete explanation of such link may be unfeasible- a more achievable goal could be the discovery of an elegant theory: the identification of a fundamental neuronal principle that may unify the rich and complex world of subjective experiences. Here I present such hypothetical principle termed neuronal reflections. It is proposed that a consensus state achieved through rapid exchange of information in local groups of nerve cells is the fundamental dynamic underlying all subjective experiences. In the visual domain, the binding of visual elements into a holistic template is achieved, subconsciously, via a hierarchical sequence of integration steps. Conscious subjective awareness emerges when the reciprocal exchange of signals between neighboring neurons in high order visual areas unites them into a unique entity. This reciprocal integration resolves the ambiguity inherent in the activity of isolated neurons- leading to a meaningful conscious image in the mind of the observer. Thus, the original meaning of the term consciousness- collective knowledge- appears to surprisingly capture an essential aspect of the neuronal dynamics leading to phenomenal experience.

INTRODUCTION

The aim of this chapter is to propose a framework for a neuronal theory of conscious awareness. Such an endeavor is tentative since the problem of conscious awareness quite likely poses the most challenging puzzle in neuroscience research. On the other hand,

*E-mail: rafi.malach@gmail.com; Tel: 972-8-934 2758/or -2441; Fax:972-8-934 4131.

subjective awareness is so immediate, personal and direct that one can not but be irresistibly drawn to understand how it emerges..

I emphasize the obvious- that at this stage any attempt to formulate a brain theory of consciousness must be very speculative due to the scarcity of the relevant experimental information available, particularly from human brain research. On the other hand proposing alternative theories, as has been done by a number of prominent thinkers in the field has its own value. First, it helps in demonstrating what kind of neuronal theories can even be considered in approaching this deeply mysterious problem. Furthermore, such frameworks can guide a research program for neuroscientists- i.e. to outline experiments that will attempt to refute a theory's predictions or improve its details.

This chapter is written with a broad audience in mind, so I made an effort to avoid any technical jargon requiring prior knowledge in neuroscience. I want also to emphasize that while I present here my personal perspective on the subject (which likely will raise many objections) much of it is not original and reflects an integration of a large body of previous ideas generated by a long list of superb scientists, philosophers and psychologists involved in this exciting field (for informative reviews see (Zeki, 2001; Rees, Kreiman and Koch, 2002; Baars, Ramsøy and Laureys, 2003; Block, 2005; Lamme, 2006; Block, 2007; Rosenthal, 2005, Tononi and Koch, 2008; Haynes 2009; Dehaene and Changeux, 2011, Searle et. al. 1997).

CAN CONSCIOUSNESS BE EXPLAINED?

Before I embark on outlining my own hypothesis, I would like to briefly address the recurring arguments that in fact scientific research is unable to contribute anything of value to the mind-body problem. It is argued that this question is so mysterious and unapproachable that it can not be envisioned, even in principle, how scientific tools can "explain" consciousness. I therefore would like to begin by discussing what can science contribute to this deep question, and furthermore, how can we even know that a certain scientific theory is successful in making progress towards understanding of such seemingly intractable problem.

My personal opinion is that the pessimistic outlook concerning the ability of science to deal with the mind-body problem is unwarranted. I think this pessimism stems from an erroneous conception of the aim of scientific theories in general. A scientific theory does not attempt to completely "explain" any natural phenomenon. If we take the theory of evolution as an example, it is clear that Darwin did not succeed and in fact has not even attempted to fully explain how new species emerge. In Darwin's time the structure of DNA was unknown so obviously major gaps still existed in Darwin's original hypothesis. The reason we consider Darwin's theory as a historical breakthrough is that his theory of evolution was *elegant*- it was able to show that a large number of seemingly unrelated observations - the physical similarity between different species, the appearance of fossils etc. could all be derived from two simple principles: mutation and selection. Remarkably, Darwin was even able to correctly estimate, using these principles, the age of the sun!

Our strong sense of understanding facing Darwin's intellectual triumph is due, then, to the theory's beautiful simplicity. More generally, a successful scientific theory is therefore not judged by its ability to fully explain a certain phenomena, but rather but its ability to

introduce elegance to a chaotic set of observations. A theory is successful if it can show how a large set of phenomena, that superficially appear unrelated, are in fact a reflection of a deeper underlying principle, provided of course that this principle is compatible with all relevant observations. This powerful aesthetic drive is really the decisive factor in the success or failure of scientific theories.

How this view concerning the nature of scientific theories applies to the feasibility of a neuronal theory of consciousness? Similar to all other scientific theories, I will argue that in this case as well, the aim of a theory of consciousness is not to explain, or “solve” the incomprehensible link between mind and brain. But rather to show how the amazingly complex world of seemingly unrelated conscious mental phenomena may be derived from a single underlying principle. This principle will have to show that the smell of chocolate, the beauty of the starry night and even our ability to reflect upon our own conscious experience- all stem from a single, simple principle. This principle will also have to define the border between conscious and subconscious mental phenomena. To distinguish, for example, between the erratic motions that our eyes make when looking at a picture – a motion which we are completely oblivious to, from the vivid sense of motion we paradoxically experience when presented with a sequence of still images during a movie.

If a scientific theory could account for all these, we would certainly consider this a major and even dramatic advance in our understanding of the link between the mind and brain. This will be true even if we will not be able to explain why it is the case that this unifying principle works.

This situation is reminiscent of another fundamental triumph of modern science- the theory of quantum-mechanics which describes the physics of phenomena at minute scales. Here the situation is far more advanced- the discovery of a set of elegant principles - described in precise mathematical equations- has in fact already been made and is justifiably considered a major triumph of science. However, we are still in the dark about the explanation (also termed the “interpretation” of these equations) - why the laws of quantum mechanics work so well. In the case of neuroscience – we are still far behind- we are not even close to a successful theory linking subjective experiences and the brain. The hypothesis I will present here both lacks in detail and in sufficient experimental support- and hence should be viewed as a tentative framework. The basis of this network is a principle that if elaborated more precisely and proven correct, could indeed constitute an important advancement towards a scientific theory of this most mysterious and ancient puzzle- how is subjective experience related to the physical brain.

THE WORD

I will begin from a vantage point that will appear at first completely misplaced. About 4000 years ago, one of the most profound cultural developments in human history occurred in the Sinai desert. This revolution was the invention of the written alphabet. It is thrilling to see how we, Hebrew readers, can still identify the letters in the first alphabetic word that has ever been written. This invention was so powerful because the alphabet is an example of a deep aspect of nature- the power of combinatorics. It reflected the realization that through different combinations of a few simple elements it is possible to generate a vast, practically unlimited,

number of unique entities. In the case of writing- this is of course the combination of different letters into words. What defines the word is the unique assembly of different letters. The individual letter, presented in isolation is not informative- but its contribution becomes critical when it is combined with other letters into a specific word. In written language each letter participates in constructing millions of different words and the number- the “space”- of potential words is astronomical. An important aspect to note is that the mere collection of letters that constitute a word is completely meaningless unless there is some process, some mechanism that “binds” the letters into a meaningful whole. In the case of letters this is of course the reader- without a reader that understands the words- the mere fact that the letters are placed near each other on paper has no meaning. In the following I would like to show how this deep phenomenon, the power inherent in binding few elements into a vast space of complex entities may in fact be the fundamental process underlying all of conscious experiences.

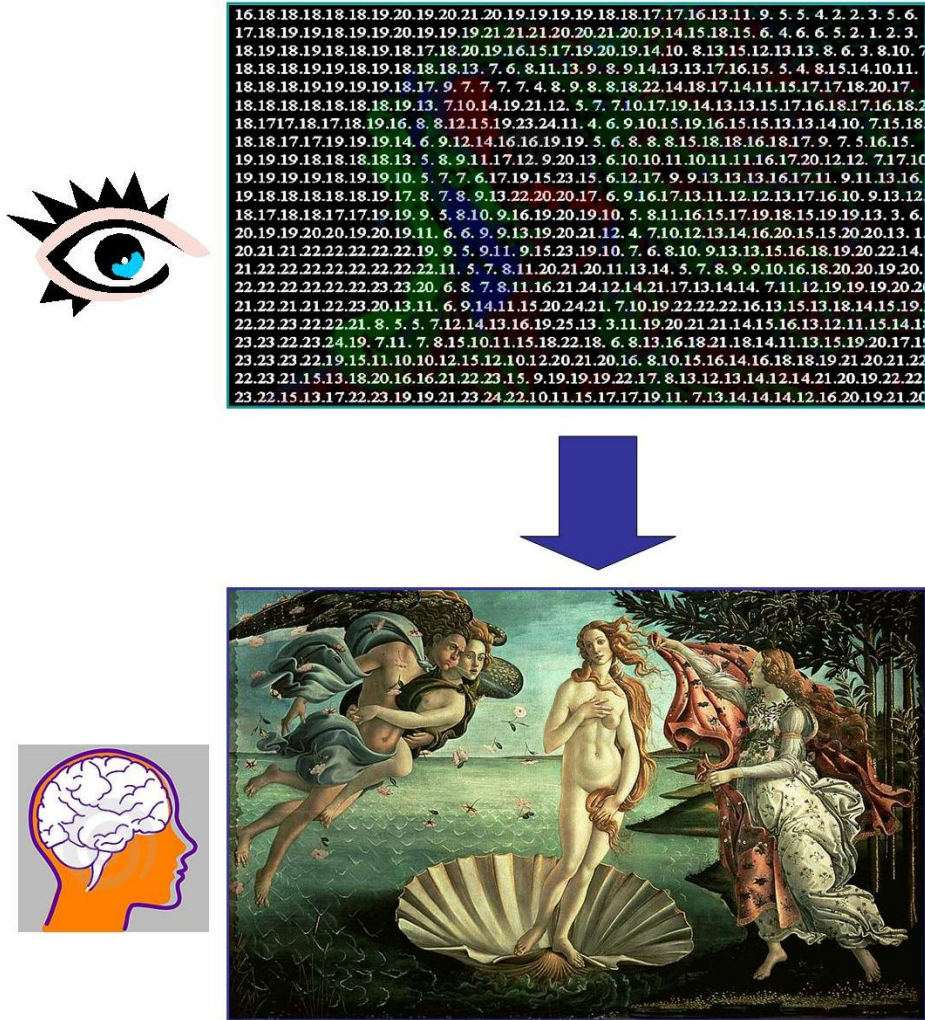
VISION AS A CREATIVE PROCESS

While my overall aim is to link brain processes to any subjective psychological experience, I will focus in this chapter on a specific subset of such experiences- the visual image. I will ask the deceptively simple question- what needs to happen in our brain so that a subjective experience of a visual image will emerge in our mind.

If we examine how the visual process begins, that is - the very moment when light particles- photons, strike the eye, we can see that at this stage there are clear parallels between the eye and a photographic camera. In both cases, photons of light are reflected from an external physical object. These photons travel through a series of optical focusing layers, both in our eyes and in the camera. They create a pattern of light- an optical pattern- displayed on a thin layer of neural tissue in the eye termed the “retina”. This tissue is actually a mosaic made of many thousands of tiny light detectors- also termed-“photoreceptors” and nerve cells that connect them. Without getting into an elaborate description of the light absorbing mechanisms of the retina, which is a fascinating topic in its own right, we can consider the photoreceptor basically as a mechanism whose role is to translate light intensity (rate of photon absorption) into an electrical signal. A critical point to understand is that at any given moment a single photoreceptor is not absorbing photons from the entire visual world but, (if the picture is sharp), it can absorb light only from a tiny point in the external world. In other words, each individual photoreceptor can report to the rest of the brain only about a single “pixel” in the external object we look at. We call this tiny pixel the “receptive field” of the specific photoreceptor. Neighboring photoreceptors have neighboring receptive fields that together tile up the entire external world- also called “visual field”- in front of us.

This simple fact, that the photoreceptors transmit to the brain a collection of isolated points has deep implications to the seeing process. To illustrate this, imagine that you are in a museum, examining the beautiful “Birth of Venus” painting by Botticelli (Figure 1). If we now ask, what information your brain actually gets from the eyes, that is- with what “raw material” the brain begins the seeing process; it is evident that the brain, in fact, does not receive any information that even remotely looks like the picture of Venus as we see it. All the brain receives is a sort of gigantic table of numbers (Figure 1) which are the hundreds of

thousands of isolated signals representing the individual pixels in the painting reaching the brain from each eye. These signals reflect the light intensities (more precisely contrast), translated into electrical pulses in each isolated point in the painting (we will not deal here with additional complexities such as wave length information, low light conditions and the fact that the information is typically delivered from both eyes). These many thousands of isolated signals are received by the brain, and in some mysterious process, that we call “seeing” our brain uses these signals to create the visual image that we experience as the breathtaking picture of Venus.



The pictures that we see are not present in the eye but are created by our brains

Figure 1.

We can summarize this as follows: the pictures that we see can not be found in the eye. In fact, they don't even resemble anything that comes from the eyes- and consequently it is wrong to say that what we see is some kind of copy or representation of the information that we receive from the external world. The visual picture that we see is the outcome of an internal, creative, process that occurs within our brains, and its relationship to the optical information we derive from the world is complex. Thus, it is meaningless to ask how the external world looks like "on its own". All sensory percepts, and visual images are no exception, are the products of neuronal activity and can exist only in the context of consciously perceiving brains.

You may protest- that this is just a philosophical word-playing and in fact we have a strong and immediate intuition that the external world looks and feels precisely as we see it. This indeed appears to be the case most of the time, and fortunately so. The fact that the visual images we "invent" are good predictors of the behavior of the physical world is what allowed us, human beings, as well as other animals, to use our vision so effectively to compete and survive in the world. However, a large number of visual phenomena- popularly called "visual illusions" repeatedly demonstrate that the coherence between what we see and what is actually out there in the world is often loose and sometimes completely incorrect.

Figure 2 provides two striking illustrations of the dissociation between what we see and what is physically there. If we look closely at the simple triangle (Figure 2, top panel) we can see that this is not an ordinary triangle. In fact it is not a triangle at all but a collection of three separated corners. The faint edges of the triangle we perceive- do not exist neither on the computer screen, nor in our retina- this is an illusion- a fabrication of our visual brain. It is important to understand the deep significance of this simple demonstration- we have here a situation in which the real physical world is made of three separate corners with no edges between them. The photoreceptors in the retina, rather faithfully inform the brain that between the three corners there is a white homogenous surface. All this is real - nonetheless, our visual brain decides, (for reasons that are not important to the point I am trying to make) to generate an image of a triangle that does not exist. In other words, in this case, its not that the brain merely fails to "represents" the external object- we see here that the brain generates an image that actually contradicts the physical reality!

An even more paradoxical illustration is depicted in the picture of the sculpture (Figure 2, bottom panel). At first sight the triangular image we perceive appears clear and quite ordinary, and we may accept it as a perfectly valid physical object. However, closer inspection of this object (the photograph is actually of a real sculpture stationed at the Weizmann Institute Science park in Rehovot Israel) reveals that such a triangle- having three right angles -is of physically impossible. Even though we fully realize that this is an object which can not possibly exist in physical reality- this impossibility does not affect in any way the perceptual clarity and ease by which we generate its visual image in our mind. Thus, we are perfectly able to create visual images of objects that *cannot possibly exist*.

The profound and deeply counterintuitive implications of this state of affairs, as has long been contemplated by philosophers from Plato and Kant onward, is that it is simply meaningless to say that the world as such looks like anything or has any visual properties- these are aspects created by each individual's brain and mind.

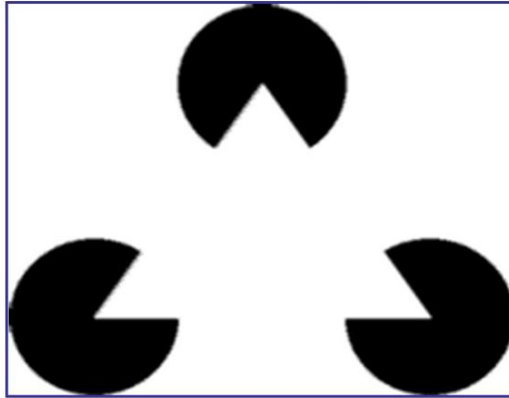
It is interesting to note the parallel of these conclusions with the picture of reality as is drawn by the physicists. Paradoxically, as modern physics advances, it becomes more unintuitive and unimaginable. Thus, one can not even begin to consider the visualization of

true physical reality- a mysterious entity whose derivative is a probability cloud spread in space and time.

In summary-despite the fact that our visual percepts of the world are so practically successful- the images we see are our own creation and they are of a fundamentally different quality from what is actually out there in the physical world.

The images we see are not "copies" of real world objects

Illusory triangle



Impossible triangle

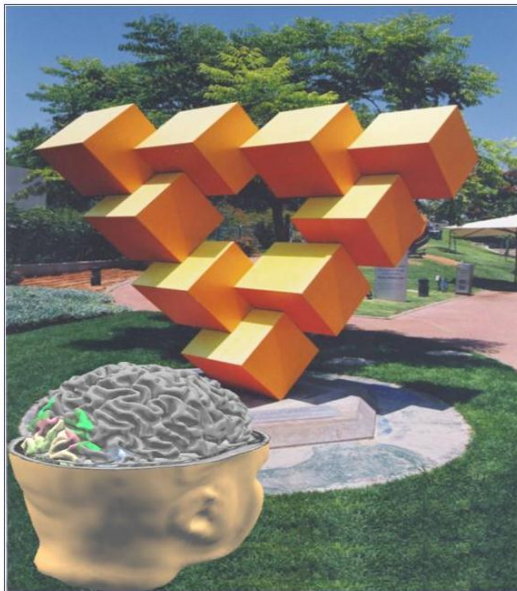
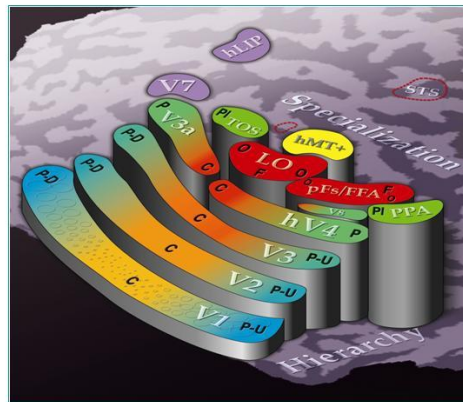


Figure 2.

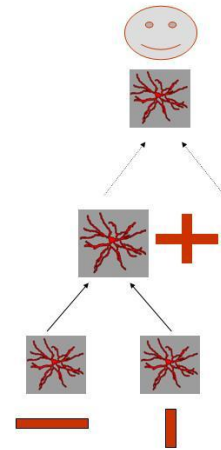
A HIERARCHICAL PRODUCTION LINE

The brain, then, is the organ that creates the pictures that we see. But how are these pictures actually generated? To understand that, we need to first consider the organization of the visual brain. Visual information, which often triggers the image formation in the brain, flows from the eyes through nerve bundles into a gate-station located roughly at the center of the brain termed the “thalamus” and from there into the outer mantle of the brain- called the “cerebral cortex”. The cerebral cortex is a sheet of tissue of about 2.5 mm thickness that covers the entire brain. In humans it is folded in a characteristic pattern (see bottom right corner of figure 2). The “visual” part of the cortex, i.e. those cortical centers that are responsible for generating the visual images, are actually located at the back part of the brain. If, due to an unfortunate accident or lesion, these parts are damaged, a person may become “cortically” blind- i.e. will lose the ability to generate a visual image even though the eyes are perfectly intact- again demonstrating that the visual cortex is crucial for the creation of perceptual images.

The Cortical Hierarchy



Atlas of human visual areas



The hierarchy principle

The elements of the visual image are integrated into a holistic template through a sequential process along the cortical hierarchy of visual areas

Figure 3.

It is important to understand that the visual cortex is not a single, homogenous, entity but is built up of a large number of specialized centers, termed also “Visual areas”. A substantial number (about 20) of different visual areas have been identified in the human visual cortex (Figure 3 left panel). However, these areas are not randomly scattered but are organized along clear principles. An important principle is the hierarchical layout of the visual areas. Metaphorically, one can envision this hierarchical principle as if the visual system constructed the perceptual image in a series of steps, along a production line. Thus, in the visual cortex, the visual signals are transferred as a sequence of messages from one visual “station” to another, and at each area the information becomes more integrated and holistic. More specifically- visual signals arriving from the eyes through the thalamus are received in the first cortical area- termed, not surprisingly, area Visual 1 (V1) or primary visual cortex, from there the information is sent to area V2, and then continues along an entire sequence of visual areas until it reaches the highest levels in this sequential “hierarchy” (see Figure 3 left panel).

What precisely happens to the visual information as it moves from one visual area to the next? In order to understand this we first need to get to know a bit better the “work horses” that are actually responsible for the visual process- these are the nerve cells -also called “neurons”.

GOSSIP MACHINES

When examining a neuron in detail, we find a remarkably complex structure. However, for the purpose of our discussion we can greatly simplify the function of a neuron to its most essential aspect. At this basic level- the neuron can be metaphorically viewed as a specialized “social center”. The essence of a neuron’s function is the gathering and transmitting of information derived from the community of other neurons. This information transfer is carried by electrical pulses that travel along the neuronal branches- a bit similar to transmitting messages along telephone wires. The signals are actually tiny “explosions” of electrical currents, which are termed “spikes” and when amplified through a speaker system actually sound like machine gun firing-hence when emitting such pulses a neuron is said to “fire spikes”. Each neuron receives such pulses from many (often thousands) of other nerve cells in a manner that can be either arousing (excitatory) or depressing (inhibitory).

A crucial aspect of a neuron’s function is the “threshold” it is the level of excitation of the neuron above which it begins to fire spikes. Thus, when the signals that impinge on a neuron bring it to a state of sufficient arousal, the neuron’s threshold will be crossed and the neuron will send a burst of spikes to its fellow neurons- the more excited the neuron will be, the faster will be the rate of its firing. Importantly, all spikes are roughly equal to each other- so the information transmitted between neurons is embedded in the timing of the spikes- mainly how fast they follow each other- i.e. in the rate by which a neuron will emit such spikes.

Metaphorically we can illustrate the essential role of a neuron as being the ultimate “gossiper”. The neuron is specialized for gathering information from its “community”. When the information is sufficiently important, the neuron will get excited and share this information with many of its neighbors. Some of these neighbors and this will turn out to be

very important, will directly or indirectly be the same neurons that sent the significant information to the neuron in the first place.

LINE DRAWING

Now that we understand a bit of the function of a nerve cell- lets consider how this “social” creature plays its role in generating a visual image. We start with the first step of visual integration, which occurs in area V1. An important discovery in this field has been made by two scientists -D. Hubel and T. Wiesel-who received the Nobel Prize for their findings. Hubel and Wiesel inserted tiny micro-wires that can capture the electrical pulses (spikes) into area V1 (in anesthetized animals) and examined what caused the neurons in this area to fire spikes. In contrast to the photoreceptors in the retina, that respond to tiny dots of light, the neurons in V1 do not get excited by such dots. Instead, what was found be particularly effective in making these neurons fire were elongated lines having a specific orientation and location in space. A neuron of this kind could be, for example, sensitive to a horizontal line in the top right corner - and will be excited every time such a line appears on the retina. But if this retinal pattern will rotate and the line will become vertical, this neuron will shut off, and another neuron, which is “specialized” for detecting a vertical line in this location, will begin firing. As a population, there are neurons in V1 that are sensitive to all orientations and all places in the retinal pattern.

How did the neurons in V1 become sensitive to oriented lines instead of tiny points? Recalling how neurons work it should not be too difficult to understand. Imagine that a neuron in V1 is connected in such a way that it receives (indirectly) information that comes only from a group of photoreceptors arranged along a horizontal line; we will call this neuron a “horizontal-line” neuron. If we could color this specific set of photoreceptors we will see that a “template” in the shape of a horizontal line will appear on the retina. Now let’s consider how the V1 neuron that receives signals from such a template will react when exposed to bright lines of different orientations. If the presented line is horizontal and its location matches perfectly the template on the retina- then all the photoreceptors that connect to the “horizontal-line” neuron will be simultaneously excited and therefore will emit signals which together will bombard the V1 neuron at high intensity. This in turn will excite the recipient neuron sufficiently to cross its threshold and hence it will start firing. However, if the orientation of the presented bright line will change, so that it will not match the retinal template, the line will now miss most of the connected photoreceptors and only a small fraction of them will be activated and send signals. This small fraction will reach the V1 neuron but will not be sufficient to excite the “horizontal –line” neuron above its threshold, hence the neuron will remain silent. As experimentalist, measuring the firing of the neuron to lines of different orientations we will conclude that the neuron is sensitive (also termed “tuned”) to a horizontal line.

If we think about the significance of this “tuning” of neurons, we can say that V1 neurons are “expecting” to see lines in particular orientations. It turns out that such a neuron can be active even if the stimulating line is not perfect- i.e. will be a bit crooked or will be missing a small part, as long as the template activation will be sufficient to cross the threshold of the neuron- it will fire. Thus, a nerve cell can to some extent “correct” and complete slight

mismatches between its a-priory expectations (which are built into its pattern of connections) and the real optical picture that actually arrives at the retina. We see therefore, that even at the earliest stages of the visual system- some creative element, some dissociation between the external optical information and the internal neuronal image already begins to appear.

PIECING THE PICTURE TOGETHER- SEQUENTIALLY

As I said above, the visual system is not a single entity but is made of a series of visual areas. In Figure 3 (left panel) these areas are depicted as a series of climbing steps. This picture emerges from many studies showing that a cortical area integrates and abstracts the information coming from the area upstream to it. We have already seen how sensitivity to an elongated line can be generated using integration of signals. Let's now consider how this process can be extended along the cortical hierarchy. We start with the fact that V1 contains neurons that are specialized ("tuned") to different line orientations. Consider now that two such neurons, one tuned to a vertical line and the other to a horizontal line send their branches to a "higher level" neuron in area V2- what will this target neuron in V2 be tuned to? Following the logic we discussed with regards to the integration in V1, we can see that by the mere convergence of signals from these two neurons and a threshold, a new kind of tuning or expectation emerges. Importantly, this type of tuning is not present in any of the inputs this neuron received from the area below it. Instead of being tuned to a vertical or a horizontal line- the V2 neuron will be sensitive to the "gestalt" of these two elements- i.e. to the symbol "+" (Figure 3 right panel). The reason for this creative ability is, as before, the threshold- each line, on its own, will not be sufficient to cross the threshold- however, when placed together in the form of a "+"- the combined excitation of the two inputs is sufficient to make the neuron start firing. While this is of course a simplified model, it illustrates how the visual system, through a sophisticated choice of connections between nerve cells can start solving the problem we considered in the beginning- how an isolated patterns of dots in the retina become a holistic picture in the mind of the observer.

In this example we used a simplified model built of only three nerve cells- imagine the richness and complexity that can be achieved if you consider thousands of such nerve cells, all interconnected and integrating information at different levels of strength and of excitation or suppression.

The hierarchical process of integration does not stop in V2 and continues through the hierarchy of visual areas all the way to extremely complex and sophisticated template tuning. Indeed, recent experimental findings from the human cortex reveal that at the top of the hierarchy the integration process is so advanced that neurons are tuned to templates that appear similar to a visual picture- e.g. for a face a tool or a topographical scene (Figure 3 right panel). These results show, for example that merely presenting elements of a face is not sufficient to excite the neurons at the top of the hierarchy. In order to get high firing the elements must be placed in the correct template configuration. Furthermore, if some of the picture elements are obstructed, yet we still perceive the entire template, the neurons' firing continues - indicating the ability of such holistic neurons at the top of the hierarchy to "complete" the missing template information.

In summary, it appears that the hierarchy principle, although simple in its basic operation- i.e. converging connections between nerve cells, actually achieves tremendous richness and sophistication in tuning neurons at the top of the cortical hierarchy. We can say that the tuning of such holistic or gestalt neurons in fact constitutes our visual memory or expectation of the visual object. These expectations are embedded in the specific structure of neuronal connections-ready to produce activation whenever the proper optical pattern appears on the retina.

THE COMBINATORIAL EXPLOSION

It appears then that the hierarchical integration of information succeeded in solving what is commonly termed the “binding” problem- the problem of combining the isolated bits and pieces of the retinal pattern into a meaningful image. Indeed we find that at the top of the cortical hierarchy single neurons appear to accomplish this binding. Could it be then that the hierarchical integration is the critical principle that underlies all subjective experiences? In other words- could it be that whenever such a “gestalt” neuron, that is tuned, say, to the template of Boticelli’s Venus, is firing- this in fact constitutes our mental image of Venus? As a historical note, this notion, of a nerve cell that tuned to a single, specific picture, has been raised, as a thought experiment, by the scientists Bralov and Letvin who coined it, tongue in cheek, “The Grandmother neuron”- arguing that perhaps such neurons exist in our brains, and if lesioned, this will prevent us from ever perceiving our grand-mother again!

Apparently we have succeeded in formulating a single principle that potentially could account for all conscious visual images- and as such indeed satisfies our search for an elegant theory. So did we finally arrive at the sought after link between brain and mind? Unfortunately, while the hierarchical integration solves the binding problem, this solution contains within it the seed of a deeper problem that makes this potentially elegant model completely unfeasible. The reason is simple- there are just too many different pictures that we are capable of seeing. If we were to allocate for each visual image that can potentially emerge in our mind even a single specialized neuron, we will quickly run out of neurons even if we dedicate our entire brain to the visual sense.

To comprehend the enormity of the space of all possible images- just consider the case of human faces. We can see clearly and likely differentiate between most human faces belonging to our own race- i.e. many millions of exemplars. Note that we are not dealing here with recognition, merely the ability of our brain to construct a clear and distinct visual image of such faces. Now consider that each of these faces could appear in thousands of different variations- with a hat and without a hat, having many different colors of hair and skin complexions, being young or old, smiling, sad or surprised and on and on. Each and every one of these permutations should be multiplied by all the faces and all other parameters leading to trillions of distinct face exemplars- and each of these unique exemplars will necessitate a special neuron! This enormous combination of different possibilities has been termed the “combinatorial explosion” problem.

It is clear therefore that we simply don’t have enough neurons in our entire brain to individually represent each visual image. And every nerve cell, even at the highest point in the cortical “hierarchy” must be tuned to many different templates. Experimental findings

from the human visual system actually support this inevitable conclusion. They indicate that each neuron, e.g. a “face” neuron is actually tuned to literally thousands or even millions of different face templates. You can metaphorically conceptualize such a neuron as a “totem pole”- i.e. a collection of different face templates stacked on top of each other. The concept is in principle similar to the notion of a “grand mother” neuron that we discussed earlier- but the totem pole neuron consists of millions of different grand mothers. Thus when we look at a face the retinal pattern will match one of the pre-existing face templates and the neuron will start firing. Because every neuron is tuned in parallel to millions of different templates, this presumably can solve the problem of the combinatorial explosion.

However this “solution” unfortunately only leads us again to a major new conundrum. Imagine that such a “totem-pole” neuron, responsible, in parallel, for the perception of thousands and perhaps even millions of different faces starts firing at a fast rate. How can we know which of the thousands of different faces in this neuron’s library is the one we are supposed to see? In other words, the activity of an individual totem-pole neuron, firing alone, is utterly ambiguous- it does not point to any specific visual image.

THE NEURAL “WORD”

Here we come back to the idea of the written word and the Alphabet that I introduced in the beginning. The beauty and power in the invention of the written word is that it was a way to unambiguously represent millions of different concepts with just a handful of basic building blocks. The idea here is that in the same way that a written word is created by the combination of different letters, so does the visual image- it is generated by the unique combination of a specific group of active nerve cells.

To see how this could happen, let’s consider again our “totem-pole” neuron, and assume it is tuned to the face of an old friend of ours- “David”. However, this is a totem pole neuron, so it is also tuned to many others faces, among them are also Doris Debby and Dov. Lets call it for convenience the D neuron. When this neuron fires, how can we tell which of all these persons we are supposed to see? The answer is that the D neuron is not acting alone: another neuron is located nearby and it is also active when we see David. However, in contrast to the D neuron, this neighbor neuron is tuned to another set of faces- in this case Mary, Jack and Larry. For convenience we will call this neighboring neuron the A neuron. Critically, as you probably noted, neuron D and neuron A are tuned to overlapping yet different libraries of faces. This is where the resemblance to the alphabet becomes apparent- just as it is with the alphabet, every neuron participates in the representation of thousands of images. However, the ambiguity is resolved through the joint activity of the neuronal group. We will use the term “cell assembly” for such a group- in honor of D. O. Hebb, a great psychologist who coined this term. Note that in our toy model, in which the cell assembly consists of just of neuron D and neuron A- it is sufficient that both these neurons will be active, for the cell assembly to resolve the ambiguity – ruling out Doris, Debby and Dov and leading to the unambiguous emergence of *David’s* face in our mind.

It is possible to experimentally estimate (very roughly) how many neurons are active when a single image is perceived and the number is likely to be no less than a million. In other words, the neuronal alphabet of the visual system consists of about a million different

letters. Imagine the vast, literally cosmological number of possible “words” i.e. visual images that can be generated in the brain with such a vast and powerful writing system.

MIND READING

For a moment it may seem that our “alphabet” model of neuronal activity finally solved the combinatorial problem- and hence the problem of conscious vision: neuronal “words” in the form of active cell assemblies are created and form the basis of the visual images that emerge in our mind. But, again, this solution is deceptive, and only leads us into an even deeper problem, which is perhaps the most mysterious and unresolved issue in consciousness research. We should recall that the meaning of a written word is not available to the word “on its own”. There is no meaning to the collection of letters “D-a-v-i-d” without a reader- some process that receives these letters and binds them into a meaningful entity,

In the brain, then, there must be some mechanism that “inspects” the activity of the individual totem-pole neurons in the assembly and recognizes this pattern of activity as representing “David”. The mere isolated activity of the nerve cells, even when they are active together, remains ambiguous- only when a process “reads-out” the pattern of activity and somehow integrates it into a holistic entity the ambiguity in the neuronal activity can be resolved. In a sense our situation has not advanced a bit beyond that of the pattern of activity in the retina- we simply substituted the isolated signals of the photo-receptors with that of highly sophisticated “totem pole” neurons. While high level neurons indeed succeeded in binding the optical elements into a holistic gestalt- as isolated nerve cells their activity remains as ambiguous and meaningless as that of a photoreceptor in the retina.

Identifying the “reader” that binds and assigns meaning to the neuronal representations in the visual system is thus a fundamental step in our search for a neuronal theory of visual awareness. Influential theories attribute such read-out function to the frontal lobes of the brain. It is known that the frontal lobes are specialized for the highest cognitive roles- such as attentional control and decision making. A particularly intriguing, although poorly understood high order function attributed to the frontal lobes is the representation of the sense of self- our ability to be aware of who we are. Although not well-defined, we do have a strong intuition that the visual image we see does not simply hangs up out there- but is actually owned by “me”, the observer. Putting it in neuronal terms- it is an attractive possibility that the same frontal brain areas that represent the sense of self- our awareness of our own being- are also the ones that are engaged in reading out and binding the fragmented pattern of neuronal activity in the visual areas into a unified conscious percept. In our concrete example- *I* am the one who sees David’s face and it therefore makes sense that frontal brain regions that are associated with my awareness of myself are the ones that integrate the isolated neuronal letters into a meaningful word and in the process endow the visual experience with its strong subjective flavor.

To examine if this intuition is correct we must rely on experiments – mainly with humans who can communicate to us such self-related mental states most accurately.

An important set of experiments that examine conscious vision are threshold experiments. In such experiments pictures are presented in such a way that they are difficult to perceive. For example, in a paradigm called “backward masking” we present a picture of a

target object for a very brief period of time, a fraction of a second, and then we obstruct this picture by replacing it with a high contrast meaningless pattern of randomly oriented lines.

Subjects sit in front of the computer screen and indicate, either vocally or through button presses whether they succeeded in seeing the target image. By manipulating the duration of the target exposure, the experimenter can bring the subjects to the threshold of their perception- i.e. arrive at an exposure duration in which the subjects sometimes see the target and sometimes fail to see it. If we now examine what happens in the brain of these subjects at such threshold experiments- we find enhanced neuronal activity in high order areas of the visual system when the subjects succeed in seeing the targets, but at the same time we often see increased activity in the frontal parts of the brain as well.

It will seem then that indeed, theories that posit a higher order “readout” mechanism in the frontal part of the brain are supported by these experimental findings. However, closer look at this problem reveals that the situation is not so clear cut. The problem is that, while we indeed see brain activation in the frontal lobes when subjects succeed in seeing pictures, the cognitive process that is associated with this frontal activity is not clear.

Alternative theories could be brought up to explain such frontal involvement. For example, it may be the case that the frontal lobes, which are known to play a role in short-term, or “working” memory, may be related an automatic tendency to memorize, at least briefly, the visual information we just consciously perceived. Even more general aspects, such as increasing our attention, or even the sense of reward a person experiences when successfully performing a difficult task – all these may produce the frontal brain activations.

Finally, and perhaps most significantly, such activity may be associated with our ability to reflect upon our own visual experience. This mental capacity, also termed “meta-cognition” is the process by which we not only experience the direct perceptual aspects of the image, but we also have the capacity to become aware of the fact that we are experiencing the percept. Note that when subjects are asked, in the course of an experiment whether they saw a target or not, they are in fact requested, to some extent, to engage such meta-cognitive processes as well.

Thus, we can see that the conventional visual research paradigm, where the scientists make every effort to control all sensory and behavioral parameters and obtain precise evaluation of the subjects’ experience, is problematic. It leads to many additional layers of auxiliary cognitive processes that may not necessarily be an integral part of the perceptual process itself.

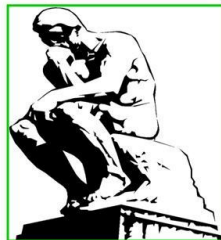
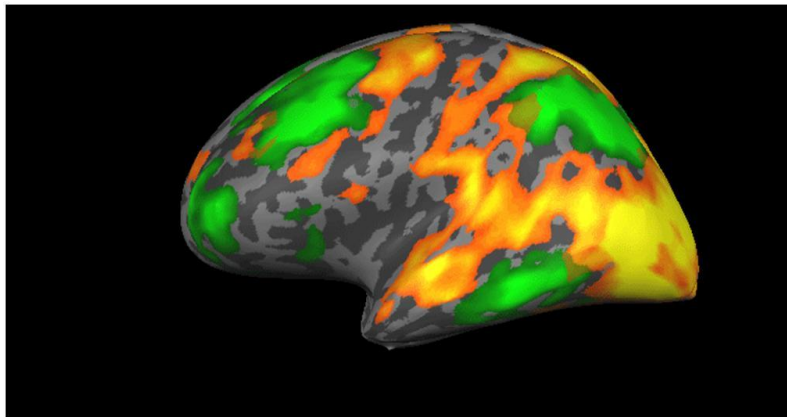
In a sense, this conundrum is reminiscent of what is called “the measurement problem” in quantum physics, where the experimental settings and the attempt to be obtain measurements that are as precise as possible actually interferes and disrupts the phenomena to be measured. In the cognitive domain our attempt to probe as accurately as we can the mental state of the subjects forces us to interfere with this very mental state and hence to distort the cognitive state we attempt to study.

It seems we are faced with a built in limitation of cognitive research that is embedded, just as in physics, in the very nature of conscious awareness. Is it then possible to circumvent these methodological limitations by studying visual experience in a bit more open and spontaneous conditions? Can we create a situation in which the subject is free to focus on the visual experience itself and is not requested to constantly reflect or introspect about it?

A Dual organization in the human brain



The Extrinsic system
(world-related)



The Intrinsic system
(self-related)

Figure 4.

GETTING ABSORBED

One psychological state that appears to resolve this “measurement problem” is a condition that can be loosely called “absorption”. This is the state you enter when you are swept away by the outside world. During absorption, we are so engaged and drawn into the immediate sensory experience that we lose all sense of self. We experience such states in moments of emergency, but often this can happen to us even during a highly engaging and captivating times such when watching a thrilling movie. In such absorbing moments we lose all sense of time and reflection- metaphorically we “lose ourselves” in the experience. During absorbing moments the subjective experience is total and immediate, and we don’t have time for introspection or meta-cognition. To study what happens in the brain during such moments,

it is necessary to simulate them inside a brain scanner. While it is obviously difficult to recreate total absorption in such enclosed and noisy environment – an absorbing state can be approximated. One particularly successful approach, we found, was to show subjects a highly engaging movie. We essentially take advantage here of the genius of great movie directors (e.g. Sergio Leone in the “Good the bad and the ugly”) to take hold of our focus of attention. Other, less colorful ways are to engage the subjects in extremely rapid and difficult sensory-motor tasks, which require their full attention and do not allow any time for more introspective interferences.

When examining what happens in the brain during such absorbing moments, we find a surprising and intriguing phenomenon. First, as expected from previous research, large parts of the cortex, particularly towards the back of the brain (more technically, the occipital, parietal and temporal lobes) including the visual system, are highly activated by the sensory stimuli. However, no less interesting than these rich patterns of sensory-driven activations is the observation that within the “sea” of sensory and motor areas, there were distinct islands that consistently *failed to* respond to the external stimulation. These were cortical areas that, instead of being excited by information coming from the external world, were tuned to some more internally oriented information (Figure 4).

While the role of each of these non-responsive islands is not fully understood yet-ongoing research is gradually revealing that indeed these regions are associated, among other functions, with self-related tasks, such as self-evaluation, voluntary decisions and autobiographical information. To reflect their self-orientation we have termed this system of areas- the “Intrinsic” system. It is also termed, for other reasons, the “Default mode” network.

THE INS AND OUTS OF THE HUMAN CORTEX

We see that using engaging movies and other highly absorbing sensory-motor tasks has, unexpectedly, uncovered a fundamental subdivision of the human cortex into two global systems. One, the “Extrinsic” system is oriented *outward*- it deals with information derived from the external world. The other, the “Intrinsic” system is oriented *inward*, and is related to information derived from the subject- the individual’s self (Figure 4).

To illustrate the complementary functional specialization of the two systems lets consider two extreme examples: consider a Zen warrior aiming an arrow, becoming one with his target. In such a complete focus on the external world, we can presume that the *Extrinsic* system is at its highest activity. In contrast, consider attending a boring lecture, when the mind “wonders” and we find ourselves detached from the external world, and instead worry about some pressing decisions we need to make. At such moments, the *Intrinsic* system is turned on while the Extrinsic system reduces its activity.

We can actually observe such swings in activity in brain imaging experiments. When subjects’ brains are scanned while they perform an intensive externally-oriented task, we see the Intrinsic system actually lowering its activity. In contrast, when we detach from the external environment and engage in thinking about ourselves-planning or recalling some personal episode, we see the swing reversing and the Intrinsic system activated while the Extrinsic system reducing its activity.

This antagonistic relationship between the two systems occurs only under extremely focused moments. Under most daily life situations both systems are co active at various degrees, reflecting the needs for combined intrinsic and extrinsic information.

However, there is one mental state which is particularly intriguing and which is worth considering separately. This is the condition called introspection- when we focus at the same time both on an external object- say a cup of coffee but at the same time we are also fully alert to the fact that we, as observing subjects, are present and are the ones who experience this cup. Interestingly, we see that under such special moments, both Extrinsic and Intrinsic systems are co-active in a fairly balanced way. It is interesting to consider that such a particularly balanced mental state constitutes the starting point of all philosophizing about the mind. This is the mental state in which we become simultaneously aware both of ourselves and our sensory experience – leading to the deep puzzlement about mind and brain.

Another fascinating question concerns the activity of the Intrinsic and Extrinsic systems during other special mental states such as meditation. It is interesting that within meditation practices, some emphasize absorption and merging with the external world- i.e. purely Extrinsic functions- for example in certain Zen practices. In other types of meditation, where the aim is a more inner focus and detachment from external disturbances, we would expect the opposite -Intrinsic - emphasis.

And in between there are the fascinating practices that aim at a perfect balance, at some merging or blur of the boundaries between the external and internal worlds. We still don't know what happens in the brain during such states- but this is experimentally tractable and constitutes a fascinating topic for future research.

How are these new insights into brain organization impact on our original problem of the “reader” and its role in conscious perception? We see that contrary to the intuitive sense we have that seeing must always involve some aspect of self-awareness- the experimental findings indicate that the opposite is true. Regions in the brain associated with accessing the self appear to be part of a specialized system that often shows *antagonistic* activity to that found in sensory areas. Thus, the self-oriented Intrinsic system appears to shut off precisely when the visual experience is particularly intense, rich and engaging. Conversely, sensory experienced is diminished when we are disrupted by self-related processes. Clearly, the Intrinsic, self-related system can not therefore serve the role of the read-out mechanism that is essential for visual perception. The principle that appears to dominate brain organization is thus the principle of specialization- in which different brain systems specialize in different cognitive functions. This specialization dominates also the relationship between visual perception and self-awareness- these two mental states will be joined and split as the situation calls for, and one can not therefore be fully dependent on the other.

ACTIVE MIRRORS

From all the above it appears that the notion that “subjectivity”- i.e. self-representation is the mechanism that “reads out” and makes sense of the neuronal activity in the visual system is actually contrary to recent experimental observations. Thus, we are back facing the main problem with which we began our enquiry- where is the “readout” mechanism- that can combine the isolated neuronal letters into meaningful holistic entities? In other words, what is

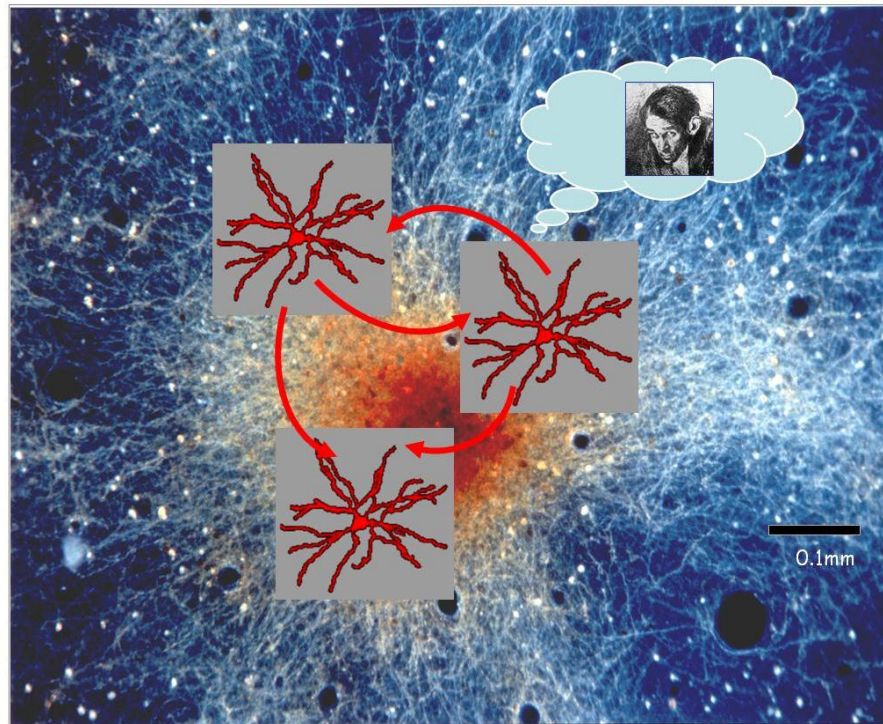
the mechanism that binds or glues together the isolated activity of the individual “totem-pole” visual neurons? As we saw above, such isolated activity is ambiguous and meaningless without readout. The binding mechanism therefore is the crucial process that endows visual meaning to the neuronal activity and thus is the key mechanisms that leads to the emergence of a conscious visual image in our mind.

I would like to propose that we should not look for the readout to occur outside of the visual system- by some yet higher order brain area. Rather, the read-out and the neuronal binding is accomplished locally by the very same high order visual neurons that form the active assembly to begin with. In other words, the same “totem pole” neurons- the neurons that constitute the individual letters of the neural “word” become a unified meaningful entity when they “read out” each other (Figure 5). To understand how such self readout can be implemented, let’s begin with the very simple notion of mirroring or reflection. But we will consider a special type of such reflection – an *active* reflection. This is a sort of a dialogue in which both the mirror and the mirrored actively take part.

It is interesting that one of the most familiar mirroring in western folklore is that of the evil queen in the fairy tale “Snow-white”. Note that the mirror in this archetypical story is not a simple mirror, of the kind we encounter in the bathroom. This is an *active* mirror, one that is both influenced by and influences back the mental state of the mirrored- the queen. In other words, among the various factors that define the mental state of the queen, we must include also the state of the mirror. On the other hand, the mental state of the mirror itself also reflects the queen’s which in turn incorporates the mirror’s and so on. We see that such a process of mutual influence (also called recursive interaction) is continuous and blurs the distinction of the mirrored from its reflection. Note that from an information point of view- that is, considering our ability to know the state of affairs, such a process of recursive interaction could be considered as binding the queen and the mirror into one informational entity. Thus, these two agents that started, before the mirroring process, as completely separate entities- in the sense that knowing the state of one had no relevance to the state of the other- now become, following the reciprocal exchange of signals between them, a joint entity- so that knowledge gained about the mirror, will also tell us something about the mirrored queen, and vice versa. Importantly, in such recursive process- we lose the ability to separate out the isolated contributions of the mirror and the mirrored- because they are entangled beyond separation through their mutual dialogue.

It is not difficult to see how this principle of mutual “mirroring” can be expanded to a group of agents. Think, for example, about a discussion group, in which each of the participants truly listens and is sensitive to the others. In such group discussion the mental state of each of the participants, can not be considered any more as isolated from the others. The state of each participant is actually a weighted sum of all the mental states of the other members in the group. This is a process of a “group mirroring” – every member in the group is not isolated any more but is both affecting and being affected reciprocally by all others- in a recursive continuous process of mutual interactions. Note that such group mirroring does not imply that the individual members in the group become identical. Each member has its own unique individuality reflected, for example, in each person’s unique sensitivity to the messages delivered by specific individuals in the group. However, through the exchange of signals, a new entity is now created from the previously isolated members of the group- a holistically unified assembly.

Local neuronal reflections



A visual image is generated whenever a group of visual Neurons Are "bound" through reciprocal exchange of signals

Figure 5.

Finally, a critical condition for group discussion to be successful is reaching a stable consensus. To see why this is important, consider a situation whereby a number of the members in the group keep vacillating, and constantly change their mind. If this happens fast enough, by the time a message from one of these vacillators is delivered to the group, they already change their mind. This leads to a mismatch between how the mental state of an individual is perceived by the group, and the true mental state of the individual. In order for all mental states to be correctly coherent with how they are perceived there must be at least a moment of stability- in which all messages have sufficient time to arrive before some members change their minds- so that all group members "agree" on what each one of them is thinking about.

With a bit of creative imagination it is hopefully possible to see the parallels between such group discussions and the exchange of signals between neurons in an assembly.

As I noted earlier, nerve cells were optimized by evolution precisely for the task of social group communication. With their numerous branches and contacts to other neurons, each cell is capable to "listen" in parallel to thousands of other group members and integrate their messages. The "mental state" of a neuron is its state of excitation which is reflected in its

spiking activity. Note that similar to the group discussion metaphor; the internal state of each neuron recursively incorporates in it the states of all other neurons connected to it, including its own prior state.

GROUP REFLECTION AND CONSCIOUS EXPERIENCE

Now, that the concept of group mirroring or reflection has been described, the central hypothesis of this chapter can be formulated: subjective experience emerges when group mirroring binds them into a meaningful entity. Put differently- the critical condition for the emergence of subjective experience is the integration of neuronal “letters” into meaningful “words” through the reciprocal exchange of signals.

Let us examine this notion in the concrete example we discussed above – the emergence of the face of “David” in the mind of an observer. We saw earlier that such a face elicits a pattern of activity in a large number of “totem-pole” neurons- neurons that are tuned to many different faces including “David”. However in order to resolve the ambiguity inherent in the multi-face sensitivity of each isolated neuron (the letters) - a group of such neurons must be bound into a single assembly or entity- in other words to become a “word”.

The suggestion is that this binding process does not depend on an external readout mechanism, but it occurs within the visual system itself. Such assembly binding happens when a number of neurons that share the same template in their library activate each other. If such mutual activation forms a stable group, this assembly unambiguously defines a unique percept, and at this moment a visual image appears in our mind (Figure 5). Note that this process is continuously recursive and can not be broken into discrete steps- thus it, in principle, could not be imitated by conventional digital computers that depend on such discrete symbols. To summarize, it is proposed that subjective experience is the product of stable recursive activation within a cell assembly.

SPLITTING THE MIND

As stressed above, any hypothesis has merit as long as it is compatible with experimental observations. Unfortunately human brain research is at an early stage of development and conclusive evidence with regards to issues of consciousness is still lacking. Thus, the notion of neuronal reflections as the basis of subjective experience, while leading to testable predictions, is only a hypothetical framework at this stage. A strong prediction is of course that preventing reciprocal interactions among neighboring neurons should abolish all subjective experience. However, we are far from being able to generate such massive disruptions, particularly in conscious humans. Nevertheless, the reflection hypothesis allows us to explain a number of known observations. I will touch here on just two of them. The first one relates to the anatomical structure of the cerebral cortex. The cortex is the main “suspect” when it comes to the generation of subjective experience, especially since, as I pointed out above, local damage to this tissue leads to immediate and often specific deficits in awareness. For example, local damage to the temporal lobes may cause a loss in the ability to see colors or to recognize individual faces.

But what makes the cerebral cortex so special compared to other brain structures? In recent years it was established that one of the unique features of the cerebral cortex is a particularly dense network of local branches inter-connecting neighboring cortical neurons. This network, called intrinsic connections (since they are intrinsic to a single cortical area), provides an optimal hardware for the type local group reflections hypothesized as critical to subjective awareness.

However, because of the way the brain is built, each cortical area is actually split in half, with each half placed separately in one brain hemisphere. Because of this anatomical quirk, at the places where a cortical area is split- the intrinsic connections need to extend long distances (from one hemisphere to the other) so as to connect neighboring cortical points. Thus, we have here a unique situation in which connections neurons that are functionally neighbors are physically “stretched” and now need to send long distance branches to interact with their “neighbors”.

An intriguing possible link between this interhemispheric network of connections and conscious awareness has been discovered in one of the most significant observations in the history of consciousness research. It was found in patients that, for clinical purposes, had the inter-hemispheric system of connections severed (to prevent spread of epileptic seizures). Detailed research conducted in such “split brain” patients revealed the astonishing observation that this operation leads to the separation of conscious awareness into two isolated sets of subjective awareness states. In such split brain patients each hemisphere became unaware of the subjective state of the other. In other words, preventing the mutual exchange of signals between neurons located across the two hemispheres led to the split of the commonly unified conscious awareness into two separate subjective experiences co-existing within a single person’s head!

How can this amazing observation be explained within the framework of neuronal reflections? In the “split brain” operation, the ability of a cell assembly to exchange signals across the two cortical hemispheres and in this manner to bind the groups of neurons into a unified entity is surgically interrupted. Because of this physical block, neuronal reflection is possible in these patients only on one or the other side of the brain and can not occur across hemispheres. According to the reflections hypothesis such separate assemblies indeed should lead to the emergence of two separate subjective states. Thus, the split brain phenomenon provides a striking demonstration of the critical role neuronal interconnections play in the formation of a unified conscious experience.

NEURONAL IGNITIONS

A second experimental observation is the repeated findings that conscious awareness is always associated with intense bursts of spikes emitted by the relevant neurons. In threshold experiments of the kind described above, we see evidence that crossing the visibility threshold is associated with a rapid and intense “ignition” within local neuronal groups. This observation points to a tight link between perceptual awareness and intense and long lasting bursts of neuronal activity. Such “ignition” is nicely compatible with neuronal reflections. Note that when neurons reciprocally excite each other- this invariably leads to an explosive process- also termed “positive feedback” - so in fact it is expected that neuronal reflections

will naturally lead to such ignitions. Furthermore, the rapid firing is critically needed for the spiking activity to successfully perform its binding function. Note that the only way in which nerve cells can reveal their functional state to others is through the signals (spikes) they emit. A substantial part of this information is coded in the length of the interval separating subsequent spikes – essentially the instantaneous firing rate of the neuron. If the neurons will use low firing rates for communication it will take too long for any unambiguous consensus to be reached by the neuronal assembly. It is as if, in our metaphorical group discussion, each person will emit one word every hour – imagine the length of time it will take to then reach a consensus. Rapid firing of neurons is thus essential for effective neuronal reflections, and if the reflection hypothesis has any merit, also to the emergence of a conscious visual percept.

SUMMARY

We certainly face a long, tortuous road in our journey towards understanding the mystery of conscious awareness. I tried to illustrate in this chapter how advancements in research methods, particularly in human brain research, opened an exciting and promising era of progress in this endeavor. The idea I try to promote is that the neuronal “group reflection”- i.e. a reciprocal information exchange in local neuronal assemblies is at the heart of conscious experience. While I emphasized here concepts and knowledge obtained from the visual system- the principle of neuronal reflections could be extended to other sensory modalities or conscious cognitive states. The difference between different subjective experiences states depends on the identity of the neurons that are united by the reflection and their relative position to the rest of the brain. It is interesting to note, in this context, that the word ‘Consciousness’ literally means “common knowledge”. It seems that the intuition of ancient thinkers anticipated modern brain research- both in the realization of the power inherent in combinatorics - as in the case of the alphabet, and in the critical role of cooperation in the emergence of conscious experience.

REFERENCES

- Baars, B. J., Ramsøy, T. Z. and Laureys, S. (2003). Brain, conscious experience and the observing self. *Trends in Neurosciences*, 26(12), 671-675.
- Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, 9, 41–89.
- Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, 30(5-6), 481-548.
- Dehaene, S. and J. P. Changeux (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, 70(2), 200-227.
- Haynes, J.-D. (2009). Decoding visual consciousness from human brain signals. *Trends in Cognitive Sciences*, 13(5), 194-202.
- Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends In Cognitive Sciences*, 10(11), 494-501.
- Rosenthal, D.M. (2005). *Consciousness and mind*. Oxford, UK: Clarendon Press

- Searle, J.R., Dennett, D.C. and Chalmers, D.J. (1997). *The mystery of consciousness*. New York: New York Review of Books
- Rees, G., Kreiman, G. and Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience*, 3(4), 261-270.
- Tononi, G. and C. Koch (2008). The neural correlates of consciousness - An update. *Year in Cognitive Neuroscience*, 1124, 239-261.
- Zeki, S. (2001). Localization and globalization in conscious vision. *Annual Review of Neuroscience*, 24, 57-86.