

Review of Zeki's “A Vision of the Brain” [5]

H. Winnemöller, M.Sc.

1 Introduction

There are many different theories of perception and not all of them are compatible with one another. This situation has occasionally been compared to the duality of light. Sometimes, light is referred to as a *particle*, at other times as a *wave*. Even though both approaches are mutually exclusive, they are very good at explaining a limited set of light phenomena. A good theory can therefore be judged by its applicability, i.e. the number of phenomena it can successfully account for. Gordon provides a set of criteria by which he judges the different theories of visual perception discussed in his book [2]. Even though by necessity limited in depth, this book gives an excellent overview of the various schools of thought behind perception. Gordon also very clearly recalls the many personal, political and religious influences, that should not, but often do, influence scientific, and especially perceptual research (other authors of science history concur). I mention this, because any scientific theory is only as good as the methods by which it was either conceived, proven, or both. With this attitude in mind, I examined related work like Hoffman's insightful read "Visual Intelligence: How we create what we see" [3]. As a follower of the German Gestalt school of Wertheimer he believes that visual perception is an active process, and not the passive one that neurologists had imagined it to be. These days, this is a widely held belief, but the Gestaltist's evidence is derived partly from relatively complex perceptual experiments and partly from thinking about the brain. Their approach is analogous to deducing the internal workings of a *black box* by controlling the input variables and observing the output responses (see Section 6 for more of this analogy). While this approach is certainly valid, the black box's function, i.e. the relationship between input and output, can never be known completely,

if only because the input domain is uncertain or too vast to be exhausted.

The approach used by Zeki [5] and others, is much more in the line of opening the black box and looking at the actual wires. The problem here is that we are not dealing with wires and electrical components, but with nerve-fibres and cells, which are far less understood. In fact, cells in themselves can be seen as little black boxes, but their input and output options are much more clearly defined. Once we comprehend the functions of these elemental cerebral components, we can start to look at the bigger picture of how they interact to constitute the wonder of the brain (see Section 2).

This approach, then, while far from complete or safe from speculative elements, yields the promise of hard, directly measurable anatomical, neurological and physical evidence, in favour of a physiologically motivated theory of perception.

The remainder of this report is structured as follows. Section 2 explains how the basic elements of the brain, the cortical cells, function and how they can be combined and interconnected to perform more and more complex functions. Section 3.2 describes how simple and complex cells are grouped into areas of uniform, specialized functions, and what these functions are. In Section 4, the different visual pathways are introduced, i.e. those conduits, that visual information travels along, from the eye to the brain. Section 5 delves into one of Zeki's most fervent arguments: the connection mechanisms of the brain, not seen as mere communication channels, but as the very essence of the brain (see Footnote 9, to see how far Zeki will go with this notion). Visual defects need to be as much part of a valid, balanced theory of perception, as the workings of the healthy brain. Section 6 therefore summarizes some of the most commonly found, or most interesting, of such defects. Zeki's own summary is quoted in Section 7, with my own conclusion following in Section 8.

2 The building block model of cells

Certain types of cells react physiologically to external stimuli by emitting electrical impulses. These stimuli may be physical (e.g. light or pressure sensitive cells), chemical (e.g. olfactory cells) or electrical (e.g. receiving input from other cells). The cells that we are interested in are those of the visual pathways. Not surprisingly, we find light-sensitive cells in the ganglion layer of the retina which is where the optic lens of the eyes focuses the image of the visual scene being observed. There are different types of light-sensitive cells situated in the retina, those that are optimally responsive to dark lighting conditions, those that operate well in light conditions and, among the latter, those who react differently

to different bands of the visual spectrum of light.

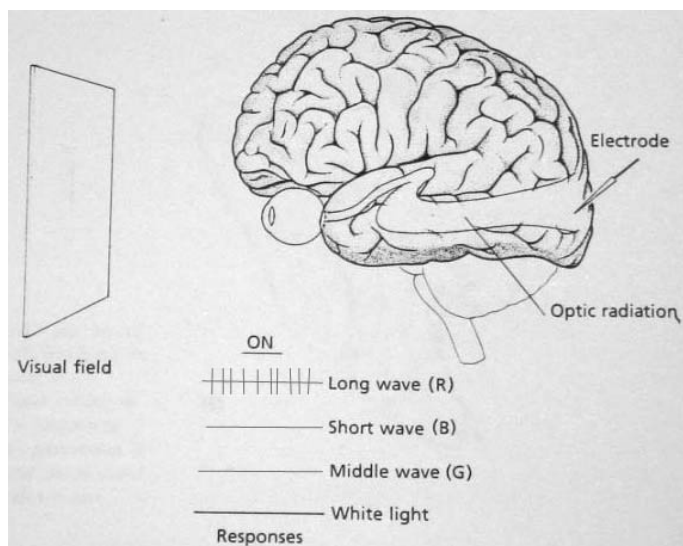


Figure 1: The response of a cell in area V4 to lights of different colours. The cell was responsive to longwave (red) light only. For simplicity, the detailed connections linking the retina to the cortical recording site are omitted. (org Fig. 14.2)

For a cell to react to a given stimulus generally means that the frequency of impulses emitted from that cell changes. An example of this is shown in Figure 1, where a cell located in area V4 (as opposed to the retina) reacts to light of long wave-length only. This requires some further explanation. It should be obvious that no light can reach the cortical areas responsible for vision directly (in fact the actual pathways that the light-induced impulses travel along are discussed in more detail in Section 4). So what does it mean that a cell in the cortical areas of the brain is sensitive to light of a particular wavelength? In essence, it means that it is connected to other cortical cells (which in turn are connected to other cortical cells, etc.) in such a way that impulses originating from retinal cells are propagated along the visual pathways to elicit a response in said cell. It should be noted that cells can receive input from more than one other cell so that many different cortical cells receive input originating, in a somewhat filtered form, from the same retinal cell. In fact, most cortical cells react to stimuli in a certain finite area of the visual field of view. This is called the cell's *receptive field* and is generally expressed in visual angles (see Figure 6).

Now, while it is still fairly easy to conceive of cells that can be *driven* (made to react) by a simple stimulus such as light of a particular wavelength, there are other cells that are maximally responsive to complex

stimuli such as oriented moving lines, or even faces. The way this can be achieved is by combining and interconnecting just the right types and number of simple cells to build up more complex cells (i.e. those that react to more complex stimuli). At the heart of this building-block mechanism are two concepts, input type and response type.

Input from one cell to another can either be *inhibitory* or *excitatory*. If several cells are connected as input to another cell, then inhibitory input from some cells can diminish or cancel the excitatory input from other cells. A classic example of such cells is the *centre-surround* cell, commonly found in the Lateral Geniculate Nucleus (LGN) and primary visual cortex (V1). Here, an excitatory stimulus in the cell's receptive field's centre is countered by the same stimulus occurring in the area surrounding that centre. It is important to note that it is not the input itself that is different, but the way it is processed by the recipient cell.

Cells can have two types of responses, the ON response and the OFF response. A cell that reacts to a stimulus *appearing* in the cell's receptive field is said to produce an ON response. A cell reacting to a stimulus *disappearing* in the cell's receptive field is said to produce an OFF response.

Similar to the primitive building blocks of a computer, these basic functions of input processing and response variation can be combined to produce very powerful *machines*. In cortical areas, we differentiate between several types of cells.

Simple cells are orientation selective, the activating stimulus being a line of appropriate orientation and falling within the correct region of the receptive field.

“*Complex cells* are orientation selective, like the simple cells, but can be distinguished from the latter by two features: they have larger receptive fields and, instead of having discrete regions from which one can obtain ON and OFF responses, both types of response can be obtained from every part of the receptive field.” (p. 78)

Hypercomplex cells are similar to simple and complex cells in that they can be driven by lines of the correct orientation. In addition, here, the line has to be of the correct length. They behave as if obtaining input from two or more complex cells, but this time using both excitatory and inhibitory inputs. E.g. if three complex cells with the same orientational preferences were to connect with one hypercomplex cell, and the input from one were to be excitatory and from the other two inhibitory, then the hypercomplex cell would respond to a line that is stopped at both ends. See Figure 2 for details.

It should be easy to see how this mechanisms could be extended to

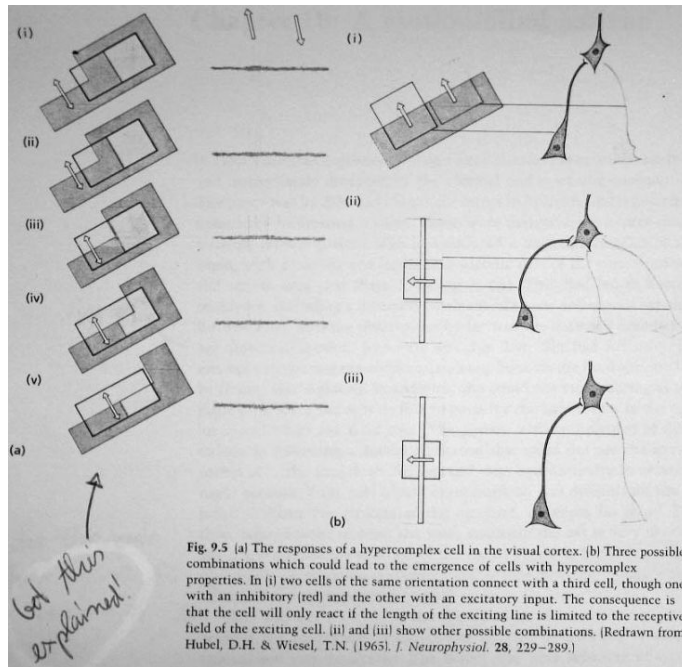


Figure 2: see original Figure caption

build ever more complex cells with more and more specific functions¹, but Zeki is careful to note that we are not dealing with a unidirectional, strictly hierarchical chain. Instead, we find that different modalities and sub-modalities of vision can be identified (see Section 3) in physically different regions of the brain and that these regions are interconnected (see Section 5) in a complex web-like structure instead of a pyramidal hierarchy.

3 Modalities of Vision and Architectural Diversity

What are the various attributes of vision? For a long time, and despite paramount clinical evidence to the contrary (see Section 6), it was believed that vision was an atomic process, taking place in a single location

¹At some stage, before the building-block model of cells became widely accepted, there existed the notion that objects in the visual world were coded by labels and that all the brain had to do in order to see was decode those labels. Playing into this idea was the concept of highly complex cells, that reacted to a very particular stimuli such as seeing a 'grandmother' or a 'fire engine' and which would be called, somewhat jokingly, grandmother-cells and fire engine-cells, respectively. While there is little doubt that cells with highly specific functions exist, we can assume that they perform some function of biological importance (e.g. detect biological motion, recognize familiar faces, etc.) as opposed to simply decoding a visual label (however that may be formed).

within the brain. The existence of a separate color centre (let alone other attributes) was heavily contested or flatly denied. How it came about that various such centres have been discovered and what these are, is explained in the following Sections.

3.1 Architecture of the brain

There are many diverse theories of vision, but most establish approximately the same visual modalities. This is reassuring in the sense that different approaches can lead to the discovery of the same underlying principle and that this principle can thus be considered to be well established. The physiological research presented by Zeki and others make this very clear and Zeki goes into fair detail to describe the techniques and processes involved.

One point that Zeki repeatedly underlines is the functional partitioning of areas of the brain and even sub-areas therein. In basically all cases a functional division coincides with an architectonic division, explained below, and it stands to reason that only technical limitations prevent the identification of those architectonic differences in the remaining cases.

Architectonic differences, referring to structural variations of different types, can be found in all parts of the brain by numerous techniques. Some of these are so obvious that they can be observed by the naked eye. For example, various distinct *shapes* of cells can be found, among these the pyramidal and star-shaped cell.

The *cytoarchitecture* of cells can be discovered by staining the ribonucleic acids within them, thus producing lighter and darker staining areas.

The nerve endings of cells are surrounded by sheaths of different thickness. These are called *myelin* sheaths and we can thus classify areas according to cells with similar thickness of these sheaths. This describes the *myeloarchitecture* of that area.

The *cytochrome oxidase method* stains cells for metabolic activity using the enzyme cytochrome oxidase. This method is said to be more specific than the cytoarchitectonic method.

Positron emission tomography (PET) is a fairly new and promising technology that is used to identify active regions in the brain during certain mental tasks. It “depends upon localizing the increase in regional cerebral blood flow during such task. The increase can be detected by injecting a radioactive substance of very short half-life, e.g. radioactive oxygen, into the bloodstream. The radioactive atoms decay and emit positrons (positively charged electrons) which travel a short distance before colliding with electrons to produce annihilation photons. These are detected by radiation detectors.” (p.136). While the immense ad-

vantage of this technique is its non-invasiveness, it has some drawbacks, chiefly its limited resolution. These are technical difficulties that will be overcome in the future.

One point that has been mentioned above is that an architectonic difference of a cerebral region almost always implies a functional difference as well. Zeki is careful to note, though, that the inverse is not necessarily true. Regions of *apparent* uniform architecture may well be found to perform different cerebral functions. The reason for this being that an architectonic difference may simply not be found using the currently available methods. It has been discovered, for example, that regions of uniform cytoarchitecture can well be sub-divisioned using the cytochrome oxidase method. It is hoped that, as more refined techniques become available, we will be able to make more and more specific architectonic distinctions, and thus improve the accuracy of the functional maps in the brain.

3.2 Modalities of Vision and Visual Cortex

Various cortical regions concerned with vision have been identified using the methods described in Section 3.1, as well as other methods and clinical reports, etc. Section 4 describes the pathways that are used to relay the visual information presented to the eyes until it reaches the primary visual cortex. In this Section, we describe how that information is further processed and forwarded to the various visual cortical regions and what their ascribed functions are.

3.2.1 V1 (Primary visual cortex) - Distributor

As detailed in Section 4, the main output from the LGN is to the *primary visual cortex* (so called, because it was thought to be the only recipient of the LGN; an observation that turned out to be false). For historical reasons and depending on the field of research, there are numerous other names associated with the same area of the brain: striate cortex (due to a stripe running through larger than usual layer 4), calcarine sulcus, calcarine cortex, and Area 17.

Many of the visual cortical areas contain topographical maps of the retina (or other areas from which they receive input). For example, neighbouring cells in the LGN are connected to neighbouring cells in the retina, including the same spatial arrangement. Like the LGN, area V1 contains a topographical representation of the retina, but it does so with a large degree of redundancy. In fact, it seems as if a topographical map exists for most of the modalities of vision. Zeki describes this as follows: “This detailed and repetitive map in V1 gives, therefore a clue as to why V1 is so extensive an area. Each small part of the field of view

a screened for one eye and then for the other eye, and each small part is simultaneously screened for different orientations, the entire process being repeated again in the adjacent millimeter for an adjacent small part of the field of view” (p.161). In fact, “V1 not only screens every given small region of visual space for one eye and then for the other eye and for all possible orientations, but for wavelengths as well” (p.174). “...In addition to everything else that it does, V1 also screens every small region of the field of view for visual motion. It is as if V1 contains a set of anatomically identifiable ‘pigeon-holes’ into which it assembles different kinds of visual signals. . . Indeed, it could be said that there are multiple maps in V1, each registering a different kind of activity in the visual field” (p.176).

Considering the connections between different visual cortical regions discussed in Section 5, we can think of V1 as a kind of *pre-processor*. It filters information into crude functional ‘pigeon-holes’ (as Zeki calls them) based on very exact retinal maps and passes them on to the specialized visual areas. Those, in turn, perform elaborate functional processing on that information based on barely identifiable retinal maps (prompting Zeki not to consider them to be retinal representations at all, but rather functional maps. Feedback mechanisms are, according to Zeki, then used to correlate a given stimulus with its exact location in visual space). Further evidence supporting this theory can be gathered by investigating the functional properties of cells in V1 and those fed by V1: “. . .the wavelength-selective cells of V1 are really concerned with the component wavelengths reflected from a surface, for example with long-wave light only or middle-wave light only, whereas the cells in V4 are concerned with the color” (p.261). “. . .physiological experiments on the motion system mirror those on the color system. Cells in area V1 are responsive to the direction of motion of a component of the stimulus (component directional selectivity), whereas there are many cells in area V5 which respond to the overall direction of the entire stimulus (pattern directional selectivity)” (p.261).

Before the very specific functions of the individual visual cortical areas were discovered it was believed that, due to the perceived uniformity of sight, vision was a more or less unified process, happening at a single location within the brain. This was called the primary visual cortex. The area surrounding it was thought to associate the percepts constructed in V1 with memories and other experiential data. The name given to this area was therefore *association cortex* (also *pre-striate cortex*). Only later was it discovered that association cortex comprises multiple, functionally and architecturally distinct, areas. Their approximate locations in the macaque monkey can be seen in Figure 3 and they are described

in further detail in the following Sections.

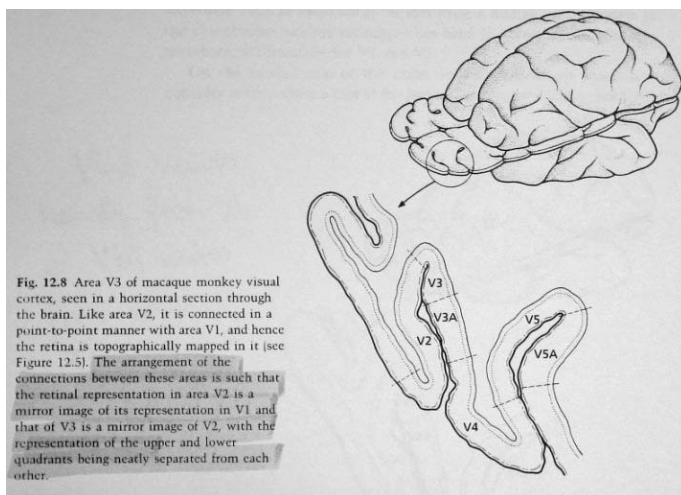


Figure 3: Location of the different visual areas within *association cortex* in the macaque monkey.

3.2.2 V2 - Distributor

From Zeki's descriptions, it appears that area V2, which immediately surrounds V1, is very similar in structure and function to V1. That is to say, it "also contains all functional groupings of cells and projects to the same specialized visual areas of the prestriate cortex as does V1, i.e. to areas V3, V4 and V5. In other words, it is subject to the same law of parallelism as V1 and, indeed, as all other cortical areas.[...] It can therefore be easily inferred that V2, like V1, will send signals related to different attributes of the visual scene to different specialized areas of the prestriate visual cortex. In other words, it acts as a segregator, just like V1." (p.177). "If sections through V2 are stained for the metabolic enzyme cytochrome oxidase, it can, like V1, be characterized by zones which are rich in cytochrome oxidase activity. [...] zones which are rich in cytochrome oxidase come in stripes of two types, thick ones and thin ones, the two sets of stripes being separated from each other by inter-stripe zones which are much poorer in metabolic activity and therefore do not stain so densely" (p.178). Because of the similarities between V1 and V2, it is interesting to look at the connections between these two areas. They "obey the 'like with like' principle. Thus, the blobs in layers 2 and 3 of V1 connect with the thin stripes of V2 and the interblobs connect with interstripes while layer 4B connects with the thick stripes." (p.178). Looking, instead, at the connections between V2 and the specialized visual areas one finds that "the thick stripes of V2 con-

nect with areas V3 [form] and V5 [(global) motion] and the thin stripes and interstripes connect with area V4 [color or form in association with color].” (p.178). One attribute of V2, which might not be defining but is nonetheless highly interesting, is the fact that it contains cells that react to ‘illusionary contours’ (p.317, see also Figure 10). These are lines that are invisible due to the fact that they are occluded by some object and can thus only be inferred. An example of this might be an automobile that is partially occluded by a street lamp. We do assume that the car’s front and back are continuously connected, even though there is a part in the middle that we cannot see directly. It is therefore enlightening that this perception is not merely an educated guess (because we have learned that most driving cars are of the connected variety), but can, indeed, be traced back to physically measurable cell activity². In fact, this phenomenon is what some Gestaltists refer to as ‘filling in the gaps’ and they describe some rules of how this is achieved (see Hoffman [3]). Seeing how this theoretical principle is backed up by physical evidence, must be highly satisfying to the researchers involved.

3.2.3 V3A and V3 - Form

Sections 3.2.1 and 4.2 discuss briefly the topographical maps that exist in area V1 and the LGN. Zeki also mentions that similar maps can be found everywhere in association cortex, but that many of these are not topographical at all. Instead, he considers them to be functional maps and assumes that the position and relation between receptive fields in these areas are related more to the function that is processed in that area and less so to the retinal position of the stimulus. It is all the more interesting to note that “the most topographic maps are those of areas V3 and V3A, the very ones which contain the highest concentration of orientation-selective cells, the ones presumed to be of importance in the perception of forms” (p.159) and “dynamic form” (p.189). Yet, it is remarkable to note that, “form perception is not the exclusive province of the inferior temporal cortex, but is a more widely distributed activity. This accounts for why there has never been a description in the clinical literature of a total and specific loss of form vision alone” (p.189, see also Section 6.2). Evidence for several form pathways (one of form alone and one of form in association with color), discussed in Section 4.4, also backs these findings.

Recent studies on macaque monkeys have shown that at least some of the orientation-selective cells in V3 and V3A are “gaze-locked, in that they will only respond to the appropriate orientation if the monkey gazes

²An attempted explanation as to how cells could react to imaginary or illusionary lines physiologically is offered in Section 5.3.

in a particular direction” (p.194). As Zeki explains, this mechanism requires some form of interaction with an egocentric coordinate system, so that the location and orientation of the observer can be related to his/her environment. It would be too far off the topic to discuss the level of self-‘awareness’ implied in these findings. Nonetheless, the fact that cellular responses can be directly linked to three-dimensional knowledge of the relation between observer and world is highly remarkable³.

3.2.4 V4 - Color

Area V4 contains a large amount of color sensitive cells. These should not be confused with the wavelength selective cells of the ganglion layer of the retina, or their counterparts in the LGN or V1. Various ingenious experiments have been performed to show this difference. One such experiment, devised by Edwin Land [4], is set up as follows. A patch of material that under normal illumination conditions (e.g. white light) would reflect mostly light of long wavelength, and thus appear red, is placed in a dark room and illuminated by three different projectors. These projectors use filters to produce short, middle and long wavelength light respectively, and their intensities can be exactly controlled. The intensities are then adjusted, so that the middle wavelength projector produces the most light, and the others only a fraction of that. The specific intensities of the projectors then define the intensity-triplet for the given set-up and can be chosen to reflect more middle wavelength light off the patch than long wavelength light. The wavelength specific cells of the retina, the LGN, V1, and V2 will now agree, that the patch reflects mostly middle wavelength (green) light, as will the color specific cells of V4, if the illuminated patch is viewed in isolation (void condition). If, on the other hand, the patch is part of a random mix of different (colored) patches⁴ and viewed as a whole (natural condition), the aforementioned wavelength specific cells still register a surplus of middle wavelength light reflected from the chosen patch, but the color selective cells of V4 correctly register for long wavelength (red) light.

This is the underlying principle of what is often referred to as *color constancy* and, in fact, represents one of the major achievements of the visual system in general: to extract constant, object-specific properties (like the ability to reflect light of different wavelengths, or *reflectance*. See Figure 4) from perpetually changing environmental properties (like

³Some more details on gaze-locked cells and related experiments can be found p.388-p.340 in Zeki’s Book

⁴The arrangement is actually not completely random, but controls for various experimental variables (such as shape, neighbouring colour patterns, etc.), so that the entire patch-work resembles the art of the old dutch master Piet Mondrian. The experiment is therefore also referred to as the *Land Mondrian experiment*.

illumination). Other such examples, like size constancy or shape constancy exist.

Returning to the abovementioned experiment, it is noteworthy that the “experiments could be repeated with [exposure to the viewer of] only a fraction of a second, with the same results, thus showing that they cannot be due to factors such as the adaptation of the photoreceptors” (p.232), which was previously believed. In fact, a similar experiment with the same set-up can be used to produce *after-images*. An after-image is what is perceived when looking at a particular shape or pattern for a while and then either closing the eyes or looking at a neutral (white) screen. The image then perceived is not blank or white, but rather the same shape or pattern just with opposing colors. This effect is the more pronounced, the longer the initial pattern was observed, but wears off quickly when looking at the neutral screen. Now, the astonishing observation, made by Land and others, is, that the after-image does not shows the opponent colors of the reflected wavelength with the highest intensity, but those of the *perceived colors*, again ruling out effects such as photoreceptor adaption.

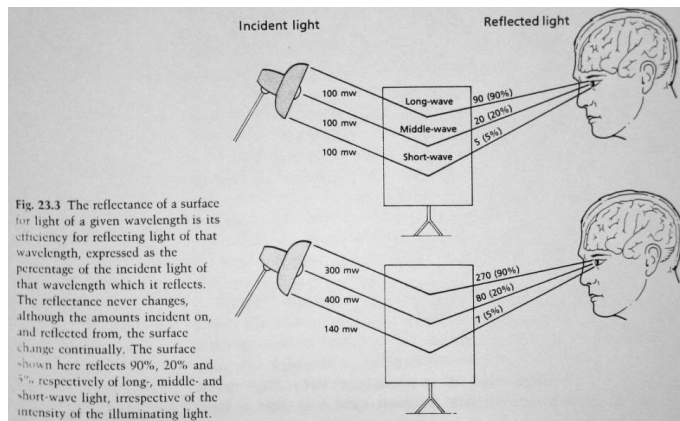


Figure 4: see original caption

The exact physiological mechanisms likely to be involved in color vision (i.e. extraction of reflectance from lightness records) are explained by Zeki in more detail than is adequate to recount here, but it should be noted that they require collection of data from a large part of the field of view. This could be one explanation for the large and topographically erratic receptive fields found in area V4.

3.2.5 V5 - (Global) Motion

When studying cells in area V5 it was found that “all cells are sensitive to motion, and over 90% are directionally selective. In other words, they

respond to motion of a visual stimulus in one direction but not in the opposite, null, direction (see Figure 5) Most cells responded to spots of light moving in one direction, although some preferred oriented lines. The interesting point was that none of the cells was concerned with the color of the stimulus” (p.123). In fact, “if one were to make a tangential electrode penetration through area V5 and record the directional preferences of the directionally selective cells at short intervals, for example of $50\mu\text{m}$, one would find a very gradual and orderly shift in the *preferred direction* of the successive cells, with only an occasional abrupt shift. By contrast, if one were to make a perpendicular penetration through the cortex of area V5, one would find that most cells prefer the same *direction of motion*, again with occasional abrupt shifts.” (pp.126).

A theme that has already been noted in various places (see Sections 3.2.1&3.2.4) is that of crude pre-computation on exact topographical maps in V1 (and V2), with the consequent exact extraction of unchanging, object-specific properties in the pre-striate areas, using crude topographical maps, or rather functional maps. As it were, the same applies to area V5. “[Some] cells in area V1 are responsive to the direction of motion of a component of the stimulus (component directional selectivity), whereas there are many cells in area V5 which respond to the overall direction of the entire stimulus (pattern directional selectivity)” (p.261). This is especially astounding as in extreme cases all parts of an object may move in contradictory directions. Again, connections (especially feedback) have to be leveraged to disambiguate this conflicting information.

Section 2 refers to the discovery that the visual system is not structured like a unidirectional, pyramidal hierarchy (which had been believed previously), but features, instead, many interconnections and feedback loops. Why this makes sense from the point of view of robustness and reliability can be observed when looking at clinical cases. For example, “lesions in V5 should cause severe disturbances in motion perception, though a rudimentary sensation through cells in V1 should remain, which has been reported” (p.302). Similar cases exist for cerebral colour-blindnesses, where different wavelengths can still be distinguished (see Section 6.1 for more details).

3.2.6 General

Other areas or sub-areas may be identified as technological advances promote such studies. Still, the findings presented so far are already astonishing and plentiful. Taking all of the above-mentioned facts into consideration, Zeki thus digests the following general points (pp. 111):

1. Several areas may reside in a cortical region of uniform cytoarchi-

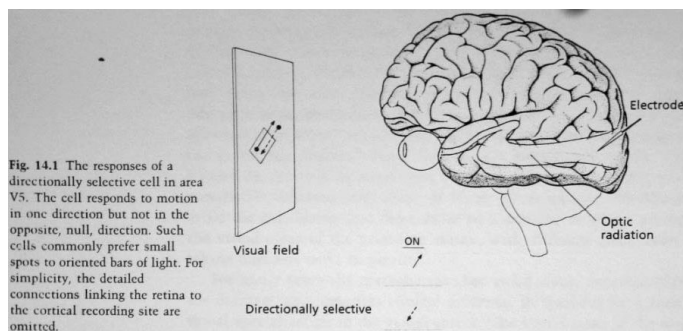


Figure 5: see original caption

texture. Consequently, cytoarchitecture may not be a good guide to the presence and extent of cortical areas. This is not to say that different cortical areas residing in a cortical region of uniform cytoarchitecture will not be found to have other architectonic differences,

2. The definition of a cortical area must rely on a number of factors. These include:
 - (a) a functional characteristic that differentiates it from other cortical areas;
 - (b) a specific set of inputs and outputs that differs from the input and output of other conical areas; and
 - (c) a unique architecture. In the case of the visual areas, another important feature is an independent representation of the retina or of the visual field, or of a visual function (see below).
3. The visual areas vary in size. If the same extent of the field of view is represented in each area, it follows that the receptive fields of cells in the smaller areas must be larger, to accommodate the same extent of representation. On the other hand, different extents of the field of view may be represented in different areas. We shall see that both conditions obtain.
4. Each of the specialized visual areas has one or more satellite areas attached to it, undertaking functions belonging to the same attribute of vision as the core area.
5. Each of the areas listed above receives an input from either V1 or V2 and commonly from both. V1 and V2 are to be distinguished

from other visual cortical areas in that all the submodalities of vision are represented in them. In other words, both contain cells which are selective for motion, orientation, wavelength and also depth.

6. Area V1 has an output to areas V2, V3, V4 and V5. The output to V4 is from the part of V1 representing the most central part of the retina only, the remaining output to V4 being channelled through V2. The output to V3A is from the parts of V1 representing the peripheral retina. Thus, the main output from V1 is mainly to the core areas; there is a far less prominent output to the satellite areas, or none at all.
7. [...] the specialized visual areas differ profoundly from each other in their functional properties. It follows that they must be receiving different kinds of signals from the primary visual cortex, area V1. It follows from this that V1 must act as a segregator, parcelling out different signals to the different specialized visual areas.
8. Each of the specialized areas, V3, V4, V5 and V6, also receives a direct input from area V2. V2 also has an output to some of the satellite areas, for example to those of the V5 complex and the V4 complex. By the same arguments given in (7.), it follows that V2, like V1, must segregate signals before parcelling them out to the specialized visual areas.

4 Light, travelling at the speed of thought - The visual pathways

This Section deals with the question of how the information reaching the eye, in the form of light, is transmitted and distributed to the different parts of the brain.

4.1 The Retina

The first (and only) cells getting into (direct) contact with light are the photoreceptor cells of the first layer of the *retina*. Two types of cells are to be found here. The *cones* are used in normal day-time vision (i.e. high light-levels and contrast). They come in three different flavours being sensitive to different bands of the visual spectrum (short, middle and long wavelengths). Then, there are the *rods*, which are used in night-time vision (i.e. low light-levels and contrast). There is only one type of rod, which explains why objects in the dark appear devoid of color.

Because of the way that each half of the field of view is processed in a separate half of the brain (discussed later), there is a natural division of the retina into left and right hemisphere. For other reasons, it makes sense to further divide the retina into upper and lower halves. We thus arrive at Figure 6, showing the retina subdivided into four quadrants.

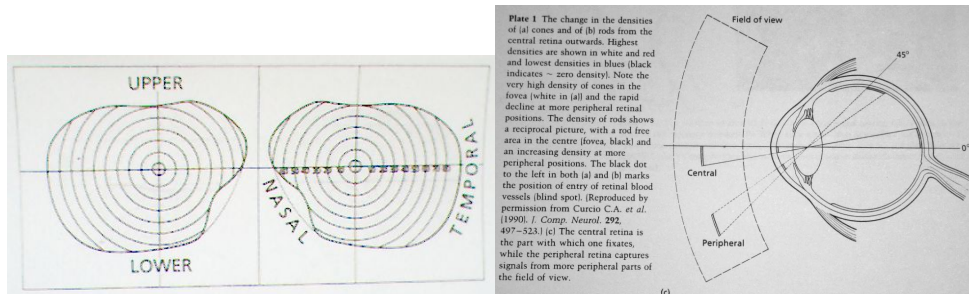


Figure 6: (Left): Subdivision of the retina. Horizontally the retina is divided into nasal and temporal halves, vertically into upper and lower halves. Furthermore, angles subtended from the focal point of the retina are also marked. (Right): see original caption.

It should be noted that due to the lens action of the eye, the nasal retina of the left eye and the temporal retina of the right eye look at the left half of the field of view (the left hemi-field) and vice versa. Similarly, the lower part of the eye looks at the upper field of view and the upper part of the eye looks at the lower field of view.

The centre of the field of view is defined to be the area of highest focus on the retina and called the *foveola*. The foveola extends over the central 1° of vision. The central 5° are called the *macula lutea*. Integrity of this part of the retina and its cortical representations are necessary for fixation and detailed vision, as well as for color vision, since the receptors for color vision - the cones - are most numerous within the central retina. Everything beyond the macula lutea is considered to be part of *peripheral vision*.

4.2 Post-retinal connections

Similarly to the first layer of the retina, the ganglion layer also consists of two types of cells. “One type, the *M cells*, is sensitive to low contrasts, responds transiently, and has axons which conduct very rapidly. It is not selective for the wavelength of the stimulus. It projects to the M layers of the *lateral geniculate nucleus* (LGN). The other type, the *P cells*, responds to high contrasts, has a sustained response and many of its constituent cells are wavelength selective. It projects to the P layers of the LGN” (p.182). Before impulses from the retinal cells reach the

LGN, they travel first along the *optic nerve* and then cross over at the *optic chiasm* in such a way that the right hemisphere looks at the left field of view and the left hemisphere looks at the right field of view, as shown in Figure 7. After passing the *optic tract*, the impulses reach the LGN.

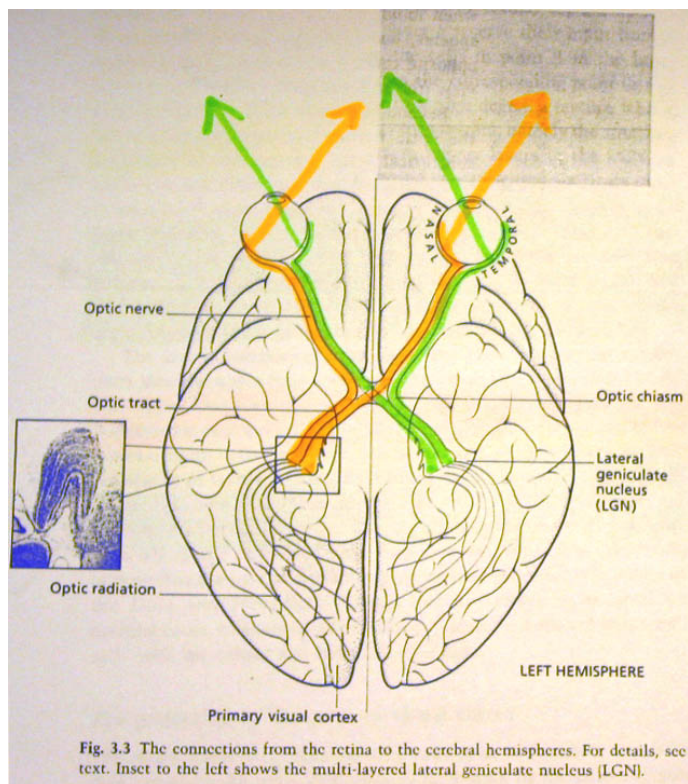


Figure 7: see original Caption text

“[The LGN itself consists of] six layers, of which the upper four have small cells and are known as the *parvocellular* or P layers and the lower two have large cells and are known as the *magnocellular* or M layers” (p.180). “Early physiological recordings from the LGN showed that most cells in the P layers have some degree of wavelength specificity, whereas those in the M layers do not” (p.180). Input from *ipsilateral* (of the same side) eye terminates in layers 5, 3, and 2, whereas input from *contralateral* (of the opposite side) eye terminates in layers 6, 4, and 1. Projection of retina occurs topographically. This means that neighboring cells in the retina project onto neighboring cells in the LGN. Topographical maps can also be found in V1, V2 and to some lesser degree in other cortical areas⁵. “Furthermore, topographical maps are

⁵The maps in the LGN and V1 are so precise, that the location of *scotoma* (blind

stacked. For example, if a point A in layer 6 receives input from a certain point in the left retina, then cells at Point B in the layer below it will receive their input from the corresponding point in the right retina.” (p.27). An interesting comment by Zeki about research attitude refers to the LGN: “[The much] studied LGN has told us very little that is of interest about vision as a process, beyond the vague statement that it may act as ‘a sharpener’ of the retinal image, a surprisingly banal function for so large and complex a structure [...] which shows that it is a mistake to avoid studying the apparently more ‘complex’ areas of the cerebral cortex simply because one has failed to understand the seemingly more ‘simple’ structures” (p.88)

4.3 Cortical Areas

The output of the LGN is relayed via the *optic radiation* (also called the *bundle of Gratiolet*) to the various cortical areas dealing with vision but mainly V1. The connections between the layers of LGN and V1 can be summarized by saying that “the output from the P layers of the LGN is relayed to layers 2 and 3, where it divides to constitute two pathways. One of these feeds the *blob cells* and is therefore concerned with color, while the other feeds the *interblob cells* and is therefore concerned with form. By contrast, the output from the M layers is relayed to layer 4B of area V1, and also divides into two components. One component feeds the orientation- plus direction-selective cells of layer 4B and is therefore concerned with motion, while the other feeds the orientation-selective cells of the same layer and is therefore concerned with form.” (p.182)

Various other connections exist, namely between the individual cortical areas and the two hemispheres of the brain. These are discussed in more detail in Section 5.

4.4 Functional Pathways

“We can now try to spin all these results together and talk of the four pathways in the visual cortex. [...] The simplest of these is the *motion pathway*. It has its origin in M ganglion cells of the retina, which relay to the M layers of the LGN. From the LGN the output is relayed to layer 4B and from there to area V5, both directly and through the thick stripes of area V1. The pathway can be referred to as an M pathway insofar as it is derived principally from the M layers of the LGN. Another M pathway, again derived from the M ganglion cells of the retina, is also relayed to layer 4B. This is the M *dynamic form pathway*, since it relays to the orientation-selective cells of layer 4B and from there to area V3,

spots of limited size) in the field of view gives a direct indication of the damaged cortical region.

both directly and through the thick stripes of V2. The P pathway has its origin in the P ganglion cells of the retina and ultimately divides to form a *color pathway* and a *form pathway linked to color*. From the P layers of the LGN, the signals are relayed to layers 2 and 3 where they feed the blobs (color) and the interblobs (form); these two subdivisions eventually relay to area V4 through the thick stripes and interstripes of area V2 and also directly.

We can therefore speak of two form pathways, one derived principally from the M system, and much more concerned with dynamic form, and the other derived principally from the P pathway and much more concerned with form in association with color. In addition to these, there is a principally M-derived motion pathway and a principally P-derived color pathway.” (p.188).

If functional representation is anything to go by, it would seem as if form is an important visual attribute. The reason for this could well be the fact that form does not only describe the shape of objects in a visual scene, but interacts heavily with other attributes of vision such as color or texture (i.e. it denotes their boundaries). The form pathway in association with color represents a strong argument in favour of this theory.

5 Interconnectednes & Feedback

Zeki, as many others before him, identifies one of the central problem of vision as this - “[The visual world which we experience is such that] one can register the precise position, shape and colour as well as the direction and speed of motion of a bus simultaneously and instantaneously, as if all the information coming from that bus had been analyzed in one place, in a fraction of a second” (p.295). This is, in fact, exactly what many scientists had believed for a long time - that because of the atomic perception of vision, vision must be an atomic process. We now know that this is not the case, and so various questions arise: ‘How can we detect and process different modalities of vision in physically different locations (see Section 3.2), yet experience a unified and instantaneous percept?’, ‘How can conflicting information, registered in different parts of the field of view, be reconciled into a meaningful, visual whole?’ (see global vs. local effects already mentioned in Sections 3.2.4 & 3.2.5). Zeki’s answer to these questions (and others, not yet posed) is *communication*. Different parts of the brain have to communicate in order to relay information. It has always been assumed that the cells in the retina, for example, have to be connected to, and communicate with, cells in the brain. Now, instead of these connections being uni-directional and strictly hierarchical, it was found that the situation is a lot more complex. Different areas

exist in the brain, which are responsible for different attributes of vision. Even though most of the input to these areas is from V1 and V2, they also receive (limited) input from each other, so that, for example, the centre associated with shape-perception, can influence the colour centre. Another, indirect, connection between the different visual areas exists via the striate cortex itself. Because most visual areas have feed-back connections back to V1, they can theoretically⁶ influence the information fed to other areas, and, indeed, that fed to themselves. The following Sections summarize the knowledge that is available about connections of parts of the brain dealing with vision.

5.1 Cerebral Connections

Many of the facts about (direct) cerebral connections with respect to vision have already been discussed. Section 3.2, for example, explains the distributor areas V1 and V2, and how they project to the specific visual areas. Section 4 describes the pathways that visual information travels along from the retina to striate and pre-striate cortex. Even though most connections from the retina pass via the LGN, it has been shown that several direct connections to mid-brain structures do exist. Other types of connections, namely those between the two halves of the brain and those that feed back into areas V1 and V2, are discussed below.

5.2 Corpus Callosum

So far, all discussion in this article refers to structures in one half of the brain only⁷, even though these structures are replicated in both hemispheres⁸. The *corpus callosum* is that part of the brain which interconnects these two hemispheres. “Due to the fact that no functional purpose had been assigned to the corpus callosum, Lashley and others supposed that its function must be to hold the two halves of the brain together” (p.292). It was later discovered that its function is, indeed, somewhat less mundane. The corpus callosum not only connects the hemispheres physically, but also physiologically. A remarkable fact is, that it does so repeatedly, once for each visual area (or function). Zeki explains this as follows: “It is not surprising to find, then, that each area

⁶As the causal nature of such events becomes very difficult to identify, the specific processes involved in feed-back are still very much under investigation.

⁷except, of course, the *optic chiasm* (Section 4.2) , which defines the crossing-over point of the optic nerve.

⁸The reason, why no distinction between the two halves has been made, is because “whether in animals or in man, it is obvious that there is no marked dominance [between the hemispheres] for the most primordial functions. There is no hint, for example, that the striate cortex of one hemisphere is dominant.” (p.165)

is connected separately with its counterpart in the opposite hemisphere, precisely because the two halves of the area (one in each hemisphere) have machineries which are different from those of other areas and thus require to be separately connected.” (p.169). This scheme, of connecting corresponding cerebral areas in the two hemispheres, is so ubiquitous, that it can almost be called a *law* - “So far, no visual area of the cerebral cortex without its own set of connections with the opposite hemisphere has been identified” (p.167). Based on this principle, researchers are now working on refining techniques to study the human corpus callosum non-invasively, in order to gain “insights into the number and organization of the visual areas in human cerebral cortex” (p.170). It should be noted, that not the entirety of each visual area is connected to its contralateral counterpart - “it is only stripes of cortex in which the midline is represented that are so connected. For example, it is only the parts of V1 in which the vertical meridian is represented that are connected with their counterparts in the opposite hemisphere” (p.166). Even though the corpus callosum thus only interlinks the limited, central part of the field of view, it’s role is essential in order to perform the kind of field of view integration that is at the core of all modalities of vision. For example, the “corpus callosum is critical for generating colour (amongst others) when the brain has to compare signals coming from the two hemi-fields.” (p.169). The same situation holds for shapes and motion that span from one hemifield into the next.

5.3 Feedback

In Section 2, receptive fields are explained and how a stimulus within a cell’s receptive field can activate that cell, thus eliciting one of several types of responses. Section 3.2.1 summarizes some of the known facts about area V1 and how it interconnects with the specialized visual areas. In particular, it is mentioned that V1 contains a very detailed topographical representation of the retina for each of the visual modalities. This stands in contrast with far less topographical maps, found in the areas projected to from V1. An important question arises, considering this kind of system: “If the topographical representation of the retina in the specialised visual areas is unpredictable, or even nonexistent, then how can the exact retinal position of a modal stimulus be identified?”. After all, what good is it to detect a moving object, if one cannot determine *where* in the field of view this object is moving?

Zeki’s theory regarding this predicament has been mentioned before. He sees the retinal maps in the specialised areas not as topographical representations of the retina, but as functional maps, structurally governed by the visual modality computed in each area. This could well

explain, why the retinal topology is relatively intact in area V3, concerned with form vision, while a topographical representation is almost undiscernable in area V4, dealing with colour (see Section 3.2 for details). Zeki also mentions that connections have been found from each of the specialized areas back to those feeding them, mainly V1. The idea is, that V1 (and to some degree V2) parcels out the relevant information to be sent to the specialised areas. These, in turn, compute a specific aspect of vision (e.g. color, form, motion, etc.) using large-scale field of view integration (including the opposite hemisphere, if necessary). To correlate an identified stimulus back to the retinal position at which it occurred, signals are sent back to area V1, with its detailed topographical maps, for exact localization. Zeki is careful to note that “this is a suggestion derived purely from anatomy and no experiments to test it have been undertaken.” (p.338). It seems as if these would be experiments well worth funding.

“Feed-back might also be important in another kind of integration, one which involves the translation of one kind of visual stimulus into another.” (p.323). “There are many other examples to show that one attribute of vision can be created from another, shape from shading and depth from motion being commonly quoted examples.” (p.324). In fact, Gibson’s *optic flow* theory, builds heavily on visual information that can be extracted from moving texture elements. Zeki describes an experiment involving shapes hidden in textures and concludes: “when different segments of the texture move in different directions, forms become immediately visible. In other words, form becomes visible, through motion. Here the brain must detect the coherent motion first, a function of area V5. The coherent motion, thus generated, must then somehow be translated to excite the same orientation-selective cells in areas V1 and V3 which [would be excited] from luminance contrast, or at least cells with a similar kind of orientation preference in the two areas.” (p.324)

That this kind of inter-modal cross-talk is not only plausible, but likely, based on anatomical evidence, is shown as follows. “The return input to V1 and V2 from the specialized visual areas is diffuse. The reciprocal input from any given specialized visual area is not segregated nor is it restricted to the territory of cells in V1 and V2 that project to that area. A good example is found in layer 4B of V1. If one were to label the cells of this layer projecting to V5, one would find them to be clustered together and separated from each other by unlabelled cells, and therefore cells which project to destinations other than V5. If one were to study the reciprocal input from V5 to layer 4B of V1, one would find that its distribution, by contrast, is not clustered and is not restricted to the territory of the cells projecting to V5, but covers

the spaces in between as well.” (p.330). “It follows that, through this reciprocal projection, V5 has the anatomical opportunity of influencing not only cells in layer 4B of V1 projecting to it, but also those projecting to other areas as well.” (p.331), but also “that the reciprocal pathways are not modular and not easily localizable.” (p.334). “In summary, the arrangement of these return connections gives us the strong impression that they may constitute one of the anatomical pathways involved, not only in integrating the visual image in the cortex, but also in generating one visual construct from another, for example form-from-motion.” (p.332)

Zeki even goes so far as to suggest that (the visual aspects of) dreams might be implemented using feedback mechanisms. “It is almost certain that [the mechanism of dreaming] must involve the simultaneous activity of several visual areas and that these areas must interact to provide the integration that is evident in dreams, since the coherent visual image is itself the product of the simultaneous activity of many separate areas. The result of this interaction must then somehow be re-entered into the cortex, as if it were coming from the outside.” (p.325). “Indeed, evidence suggests that during dreaming there is a massive increase in cerebral blood flow not only in the visual areas of the prestriate cortex, but also in the striate cortex itself, and probably in area V2 from which it is difficult to separate in the low-resolution positron emission tomography (PET) scans” (p.326). Unfortunately, Zeki does not describe, what actually *drives* dreams, i.e. if the stimulus does not originate from the real, external world, then where does it come from? The severity of this question is compounded by the fact that Zeki insists that no proof has ever been found of a *Master Cortical Area*⁹ (i.e. one, which ultimately *controls* the brain and which all other areas *report* to). This means that there is, most likely, no ‘puppet master’, that could orchestrate a given dream and control the individual visual areas to produce the necessary phantom images. In addition, it has been argued that the specialised visual areas require some feed-back connections to properly localize a given stimulus, but Zeki implies that this feedback mechanism is powerful enough to invertly *drive* the primary visual cortex and thus generate cell responses at the correct topographical position. As this possibility of interaction is not mentioned anywhere else, this added

⁹A valid proof would be, for example, if an area were to be found, which has only incoming connections, but no outgoing ones. This would then, in conclusion, be the center that other centers report to (going further, it would probably also be the seat of the *soul* or *consciousness*). No such area has ever been identified. Indeed, it seems a ubiquitous property of the brain that all areas are connected to at least one other area, but generally more.

complexity would surely have to be justified by more vital processes than mere dreams? Zeki's dream theory seems well motivated, but is not accompanied by enough hard evidence to account for more than a personal conviction.

Lastly, some visual disturbances could be understood in terms of damaged feed-back mechanisms (see Section 6 for details).

5.4 Synchrony

Another important question arises from the perception model described by Zeki in terms of parallelized large scale integration: "If different cells with non-overlapping receptive fields are stimulated, how do we know if they are reacting to the same stimulus or different stimuli?". "[This] is commonly referred to as the *binding problem*, a critical problem for visual physiology." (own italics, p.321). "Suppose that three cells in area V3, all responsive to the horizontal orientation but with adjacent rather than overlapping receptive fields, are activated by the same horizontal stimulus, for example the upper edge of a fence gate. Suppose further that these three cells receive inputs from twelve cells in area V1 with a corresponding orientation which, in turn, are responding to the same stimulus. The task here is to ascertain that the three cells in V3 and the twelve cells of V1 are all responding to the same, and not to different, stimuli." (p.321). This problem, and a possible solution, are illustrated in Figure 8.

But the real issue is even more complex - "There must be some signal, some indication, that the *colour* signalled by cells in one area belongs to the same object as the *shape* discriminated by those of another and the *movement* indicated by a third. There is, in brief, the binding problem writ large." (own italics, p.323). Zeki believes that the solution is *temporal synchrony*, i.e. the simultaneous firing of cells responding to the same stimulus. "This naturally raises the question of how the cells responding to the same object in the same, as well as in different, areas come to fire in temporal synchrony. It is a subject about which we know very little in physiological terms." (p.323). This is a difficult matter indeed, for *who* or *what* decides whether two event occurred at the same time or not (especially when considering that Zeki opposes a Master Cortical Area. See Section 5.3 and Footnote 9)? Zeki addresses this as follows: "Note that the synchrony itself will demand feedback, for otherwise we are faced with the same old problem of who it is that determines that the cells are firing synchronously." (p.349). Points 4 to 7 of Zeki's concluding summary (see Section 7) deal with this issue briefly, but are not repeated here.

Suprisingly, one very much related issue is not discussed at all in this

light. Motion or movement is, by its very definition, an event that occurs over a finite period of time. Those cells in V5 that respond to (global) motion (see Section 3.2.5) must therefore have some kind of knowledge about the past and present visual configuration of a scene in order to construct motion. If the synchrony argument holds, and there is some initial, though not substantial, physiological evidence to support it, then wouldn't it seem feasible that the same, or a similar, mechanism is used in the detection of motion? Since no indication is given as to how the motion-selective cells of the visual cortex actually work, this question remains unanswered¹⁰.

5.5 Plasticity

Connectivity on all levels seems key in a healthily, functioning brain. It is therefore instructive to ask the questions “When and how are these connections formed and are they alterable or completely rigid?”. “A totally hardwired nervous system might make sense in a rigidly coded world, where everything can be identified by a unique label, where nothing changes and where every condition that an organism is likely to encounter, and every possible reaction to it, is predetermined and known beforehand. This, of course, is far from true [...]. One should expect to see, therefore, a certain amount of plasticity or modifiability in cortical connections” (p.207).

Are the cortical connections present at birth, or are they formed at some (early) stage in life? This is, in fact, the age-old problem of *nature vs. nurture*. Studies show that “when born, monkeys come with what appears to be an intact visual apparatus from retina to cortex, at least as judged macroscopically. Recordings from the visual cortex of newborn monkeys show that some cells at least are orientation selective and that in monkeys which are only a few weeks old the ocular dominance columns are remarkably similar to those found in the adult monkey, suggesting that these properties are to a large extent genetically determined.” (p.217). “Vision [to some degree] is genetically predetermined, and [...] has to be nourished or ‘educated’ during an early period of life.” (p.216). Deprivation-experiments have been performed on cats and monkeys with the result that “stimulation through the eye that had been closed during the critical period was completely ineffective in driving the

¹⁰The existence of cells responding to motion (usually in a preferred direction) is unquestioned, yet their modus operandi is much less obvious than that of, say, the colour or orientation selective cells. Wavelengths and the spatial arrangements of stimuli are instantaneously measurable, tangible quantities. Motion, on the other hand, needs to be evaluated over a period of time. The fact that they are fairly simple cells, not complicated chronometers, that react to motion in their receptive field, makes their operation so much more astounding.

cells.” (p.218). Later experiments proved that a complete deprivation wasn’t even necessary to severely and adversely affect the visual system. “The eye is disused if it is not properly stimulated with *patterned light*. The presence of light, as such, was not in itself sufficient. Thus, fitting the eyes with translucent lenses, allowing light to pass through but depriving the animals of the ability to see forms, had equally severe consequences, the cells again being unresponsive to visual stimulation [...]” (own italics, p.218). “At the cellular level, there is a critical period, during which adequate visual stimulation is mandatory if the animal is to be able to see at all” (p.218). It is interesting to note that, for a given species, this critical period is very well defined and fairly limited (in the region of a given week after birth). Deprivation after this period inflicts no lasting deficiencies on the subject.

Another related question is this - “How is the brain affected once the connections have been established and are subsequently damaged?”. Because parts of the brain can easily be injured during an accident or through a stroke, examples of patients with severe, but not complete, brain damage are no rarity (see Section 6). Associated defects are naturally not limited to vision. “A lesion sustained in adulthood and affecting the speech areas of the brain impairs speech permanently, but an individual who sustains a similar lesion in early childhood can nevertheless learn to speak [...]. Child and adult, alike, become speechless after such an injury, but the child will speak again, and does so, normally, after a period of months. The adult may or may not do so, depending upon the severity of the injury” (p.220)

”But what, in terms of hard currency, that is, the physiological properties of cells and their connections, does plasticity mean?” (p.221). Experiments related to the ones mentioned above showed that “[the area responding to] the open eye had expanded its cortical territory and taken over cortical space which, by genetic right, had belonged to the other (closed) eye.” (p.221). This is not to say that any physical growing of cells took place, but rather that cells that should genetically have reacted to the closed eye were unstimulated by the appropriate input cells and thus re-connected to perform a useful function. “Here we begin to obtain some insight into plasticity. The connections established between eye and brain are not rigid. They are labile. They can expand to take over territory not belonging to them genetically, but their ability to do so declines with time. Moreover, reverse suture experiments, in which one eye is sutured for a few days, followed by closure of the other eye, during the critical period, have shown that this process of encroachment, of the taking over of territory belonging to the ‘bad’ eye, is not an irreversible one, especially during the critical period.” (p.222).

Zeki believes that an active competition for cortical territory exists and that only cells receiving proper stimulation can fend for their space. “If the axons belonging to one eye do not transmit signals to the cortex, they would be at a competitive disadvantage compared with the axons belonging to the other eye, which will consequently monopolize more cortical cells and eventually more cortical space.”¹¹ (p.223). “Plasticity, then, is the manifestation of the ability of cells to compete for space on a third, common, cell.” (p.223).

“Experiments suggest that the plasticity of the adult brain may be much more extensive than has been supposed. In these experiments, the sensory nerves coming from the forelimbs were severed and the map in the cortex was studied twelve years after severance. It turned out that the representation of the lower face area had expanded into cortical territory that had belonged to the arm and hand, a distance of between 10 and 14mm.” (pp.225). Yet, the plasticity does not appear to be unbounded. In another amputation experiment “it was found that the cortical territory occupied by the digit before amputation was occupied by other digits following the amputation [...]. The fact that such expansion was limited to 600 μm , leaving areas of silent cortex, would suggest that the field of action, or the cortical territory, that a healthy digit can claim is not unlimited, for reasons which we do not understand.” (p.224).

6 Visual Defects and their explanations

Several theories of perception exist, which do not rely on any anatomical or neurological findings, chiefly Gestalt theory and Empiricism. By way of gross simplification, their main approach is to see the brain as a *black box*. The internal wiring is unknown so that only through clever manipulation of the input and detailed observation of the output, the functioning of the box can be deduced. Even though cumbersome and lengthy, this approach can yield remarkable results. Another vital source

¹¹This was tested in an ingenious anatomical experiment: “There is a segment of the LGN, known as the monocular crescent, which represents a part of the extreme periphery of the nasal retina for which there is no corresponding temporal retina [...]. The cells here project to the corresponding monocular crescent of the cortex and there cannot be any competition for space between the two eyes. If the shrinkage and atrophy of cells in the LGN and the loss of cortical territory are to be explained by competition, then it follows that the cells of the monocular crescent should not atrophy when the eye is closed during the critical period, for the simple reason that they have no competition even though they are not able to transmit signals to the striate cortex. By contrast, the cells in the remaining part of the same layer, those which have to compete for cortical space with the input from the normal eye, will shrink and atrophy. This is precisely the result that was obtained.” (p.223)

of information are anomalies. If all black boxes are more or less the same, but some exhibit a small defect, then the latter are likely to react differently to the same input. If, for example, a person is injured and, as a result, cortically color blind (i.e. the retina is unaffected, see achromatopsia below), but otherwise unimpaired, then we can deduce that: a) color is a function of the brain; and b) it can be affected without affecting other functions¹². This Section then deals with visual defects and how they impinge on visual perception.

Depending on the extent of a cortical lesion, the visual area affected may vary (see *receptive fields* and *topographical maps* in Section 2). If, for example, only one half of the field of view is completely compromised, we are dealing with *hemianopia*. Similarly, *quadrantonopia* only affects the upper or lower quarter of one hemi-field. *Scotomas* are isolated blind spots in the field of view.

The lesioned area is also important with regards to the symptoms to be expected. As V1 and V2 act as distributors of signals, their functioning is of vital importance to the entire visual system. In cases where these two areas are damaged, knowledge of the visual world can almost never be acquired. This is to say that some rudimentary stimulation may occur (through the minimal direct connections to the specialized visual areas), but that feedback mechanisms and integration are so severely compromised that patients do not *know* what they *see*. Inversely, if V1 and V2 are healthy, but a specialized visual area is affected, then a modality is generally not completely compromised. If, for example, V4 is damaged, then the patient will not be able to perceive the phenomenon color, but may well be able to distinguish wavelengths through the wavelength-selective cells of the primary visual cortex. On the other hand, and again due to compromised feed-back, it is likely that such a patient will also suffer from other, related, complaints, since the form-from-colour mechanism will be similarly affected.

Some of the most common visual defects are briefly introduced below.

6.1 Color related defects

There are various conditions in which patients suffer from some form of color related deficiency. The most common one is retinal in nature and genetically passed on between generations (recessive in females). The most common symptoms for retinal color deficiencies are, a shift in the wavelength sensitivity of cones, and degenerate or missing cones. Most often only one type of cone is affected (*dichromasy*), so that total color

¹²For brevity, this example is somewhat contrived. In most clinical cases, more than one area of the brain is affected, even if only because the brain is so heavily interconnected (see Section 5).

blindness (*monochromasy*) is extremely rare. Dichromasy can in turn be divided into *protanopia* (affecting long-wavelength, or "red" cones), *deutanopia* (affecting middle-wavelength, or "green" cones), and *tritanopia* (affecting short-wavelength, or "blue" cones).

If, on the other hand, the retina is intact and the problem is situated in cerebral cortex, we speak of cerebral *achromatopsia*. Taking up the example introduced in the beginning of Section 6, where area V4 is severely compromised while V1 is more or less intact, it has been recorded for a patient that he "could discriminate wavelength differences of 10-20 nm, although he described them all in terms of shades of grey [. . .]. Patients appear to be able to detect differences in wavelength but cannot use the signals to construct colors. Their knowledge, in brief, is limited to such capacities as the wavelength-selective cells of V1 and V2 have" (p.310).

Due to the interconnections and feed-back mechanisms discussed in Section 5, it is commonly observed that a lesion of one part of the brain will also affect physically removed, but functionally related areas. This holds true particularly for color. "[. . .] it is difficult to imagine an area dealing with color which is not at the same time concerned with form, or at least with boundaries, since boundaries are critical in the generation of color. Moreover, every object or form in our field of view has a color, be that color only a grey. As well, all colours, being confined in space, have a form. The two, form and color, are therefore not easy to separate." (p.269). Yet, we find that "Achromatopsia is often *not* accompanied by severe degradation of form vision. This is strange because V4 comprises both wavelength and orientation selective cells (though the latter are also wavelength specific to some degree). One would therefore expect form vision to be compromised as well, assuming that orientation-selective cells are somehow involved in form vision and that human visual cortex is similar to that of the monkey, at least at this level." (own italics, p.269)

"[Achromatopsia] has an obverse, a somewhat surprising condition in which color vision is relatively spared while the other attributes of vision are severely compromised." (p.267). These symptoms are often associated with carbon monoxide poisoning. An account of such a patient is described in Section 6.2.

There exists another clinical condition, called *color anomia*, in which the patient can not only distinguish wavelengths, but also colors, yet is unable to name those colors.

6.2 Form related defects

“Patients who suffer from form vision defects can never be said to have a total loss of form vision [...]. They are not therefore ‘form blind’ in the sense that an achromatopsic patient is ‘color blind’ [...]. One reason for this may lie in the fact that orientation-selective cells, presumed to be the basis of form vision, are so ubiquitously distributed in the visual areas of the prestriate cortex. In particular, one finds a massive concentration of them in areas V1, V2, V3 and V3A and a scattering in other areas.” (p.270). Patients impaired by form related syndromes are said to suffer from visual form agnosia (or *visual agnosia* or *object agnosia* for short). In most cases the problem lies not in the detection of edges or boundaries, but in the integration of form elements into groups and objects. One such patient “has intact registration of form elements (single lines and edges), but ... his ability to integrate these elements into “perceptual wholes” is in some way impaired. The intact information about the local form elements enables him to make accurate copies of stimuli he cannot identify.” (p.315). Figure 9 was drawn by that patient with astonishing detail, yet he was unaware of the scene he copied.

As mentioned in Section 6.1, the effects of carbon monoxide poisoning are often disastrous to most of the visual system, except color. An account, of a patient thus afflicted, reads “When the patient appeared in hospital she was totally blind. But after two days she could distinguish white from black. After two weeks she recognized all colors and ‘was able to name correctly the colors associated with various objects or the objects which corresponded to colors and could match various shades of the same color. She was able to match colors according to their intensity’. Her ability to recognize pictures and objects remained defective, the patient having ‘to recognize by adding up parts instead of by simultaneous perception of all parts.” (p.276). Another patient “described her experiences in very similar terms. Shown a green battleship, she mistook it first for a fountain pen, then for a green knife before identifying it as ‘a boat’. She explained, ‘At first I saw the front part. It looked like a fountain pen because it was shaped like a fountain pen. Then it looked like a knife because it was so sharp, but I thought it could not be a knife because it was green. Then I saw the spokes and that it was shaped like a boat, like in a movie where I had seen boats. It had too many spokes to be a knife or a fountain pen’ [...]. These descriptions are so representative that they apply to most agnosic patients.” (p.316). Two points are very interesting to note here. Firstly, individual attributes of objects can well be identified (e.g. shape, color, etc.), but not automatically integrated into a meaningful whole. Secondly, this integration, which usually happens unconsciously and effortlessly, can be

consciously emulated using the partial information available. It is evident that a conscious effort, involving powers of reasoning, experience and logic, requires far more time and results in a significantly inferior perception, than the natural, automatic process¹³.

A closely related condition, called *simultagnosia*, is the “inability to perceive more than one object in the field of view at a time.” (p.315)

Interestingly, but maybe not surprisingly, object agnosia can also affect the perception of non-figures, so-called *illusory figures*, like the one depicted in Figure 10. “[One] patient, . . . was not able to ‘fill in’ perceptually the gaps in the Kanizsa triangle [. . .] presumably because ‘a patient using unintegrated information about form may. . . faithfully reproduce a gap in the figure because the information about the overall form is not available to “drive” the filling-in process’” (pp.316).

That ‘what you see is not always what you get’ is illustrated by the following account: “[The patient] showed ‘poor perception of shape or orientation, whether this information was conveyed by color, intensity, stereopsis, motion, proximity, continuity or similarity’. In spite of this, she was able to reach for the visual objects very precisely [. . .] This dissociation suggests that at some level in normal brains the visual processing underlying “conscious” perceptual judgements must operate separately from that underlying the “automatic” visuomotor guidance of skilled actions of the hand and limb” (p.350)

6.3 Motion related defects

One fairly bizarre condition, in which the patient is unable to perceive objects in motion (usually linked with other symptoms), is called *akinetopsia*. For one forty-three year old female patient¹⁴ “[various lesions to non-striate cortex] had lead to several problems, including a difficulty with calculations and a mild aphasia [a speech-related defect]. But the most striking observation by far was the patient’s inability to see objects in motion. So severe was this that, ‘She had difficulty, for example, in pouring tea or coffee into a cup because the fluid appeared to be frozen, like a glacier. In addition, she could not stop pouring at the right time since she was unable to perceive the movement in the cup (or a pot) when the fluid rose. The patient also complained of difficulties in following a dialogue because she could not see the movements of . . . the

¹³The same or similar cases are leveraged by Hoffman and others to support their respective theories.

¹⁴As this was the first reported case of these symptoms (or at least the first taken seriously), the patient has claimed something of a celebrity status, if you will call it that, amongst perceptionists. Since any theory of perception should answer to the problems of defected vision, her’s is an often-cited case-history.

mouth of the speaker'. She had difficulty in crossing roads because of the cars whose exact position was difficult for her to judge – ‘When I’m looking at the car at first, it seems far away. But then, when I want to cross the road, suddenly the car is very near’ – she could not see the position of the car in between. In brief, she had no knowledge of visual movement and little knowledge of the visual world when it was set in motion.” (p.82).

Similar to the carbon monoxide poisoning cases described above, where only the attribute of color is relatively spared, there are cases where the basic perception of motion remains unscathed. Riddoch had examined soldiers injured by gunshot wounds during the Great War and found that “patients could commonly see movements in their blind (scotomatous) fields, though they were unable to appreciate the other attributes of the moving stimulus” (p.83). As the patients are reported to actually *see* (as in perceive) the motion, and the affected area is restricted to scotoma, this description is probably different from the symptoms of blindsight, discussed in Section 7.

An interesting effect, commonly referred to as ‘Motion without Movement’¹⁵ is associated with the non-defective brain, and can thus be observed by most people. Zeki believes that “the perception of motion in a figure containing no moving component is the consequence of the activity in a specific area of the brain” (p.280) that is stimulated by that figure. He even went as far as performing PET scans on subjects viewing such a figure, and found that an area overlapping V5, but not identical to it, was activated.

6.4 Other defects

One illness that Zeki describes in some detail is somewhat related to agnosia, yet at a higher, and more specific level. It is called *prosopagnosia* and refers to the inability to recognize familiar faces. This is strange insofar, as many patients suffering from prosopagnosia are quite able to detect form and even facial features, but, as in agnosia, they are unable to correlate the individual pieces meaningfully. One patient recounts his experience like this “I can see the eyes, nose and mouth quite clearly but they just don’t add up. They all seem chalked in, like on a blackboard . . . I can also see enough to decide whether my now-limited hair is in need of a brush, but at the same time as this I don’t seem to have enough detail to know whether. . . my face is dirty” (p.327). “There are many interesting accounts of this, even occasions when a prosopagnosic patient could see but could not recognise his own face in a mirror.” (p.327). Even

¹⁵An account on how this phenomenon can be exploited in Computer Graphics is given in [1]

more remarkable is the fact that, such patients can commonly recognize facial expressions. “This implies that the recognition of familiar faces and the recognition of facial expressions are two different things which, because of functional specialisation in the brain, depend on two different neural systems, each of which can be specifically compromised.” (p.305). “The demonstration that the recognition of a face as a familiar face can be compromised independently of the recognition of the expression on that face, or that the recognition of a person by his gait can also be specifically compromised, suggests that the memory systems are themselves highly specific and that it is these specific systems that have to be re-entered onto the constructed image.” (p.329). “In the absence of that re-entry, the face is seen as a face and is understood to be a face, but is not seen as a particular face, nor is it recognized as a particular face, because the re-entrant neural mechanism required to integrate the incoming signals with the memory trace is defective.” (p.328). Unfortunately, Zeki does not clarify, what exactly a *memory trace* is, or how it would integrate in the healthy subject¹⁶.

Zeki also mentions an account of ‘depth blindness’ (p.282), even though this should be just as unlikely as total form agnosia, due to the fact that depth-information can be extracted from so many different clues, including stereopsis, shading, texture, occlusion, etc., which should therefore be redundantly available.

7 Zeki’s Summary

Zeki himself summarizes his findings as follows (pp.355):

1. The brain strives to acquire a knowledge about the permanent, invariant and unchanging properties of objects and surfaces in our visual world. But the acquisition of that knowledge is no easy matter because the visual world is in a continual state of change. Thus, the brain can only acquire knowledge about the invariant properties of objects and surfaces if it is able to discard the continually changing information reaching it from the visual environment.
2. Functional specialization is the strategy developed by the brain to acquire a knowledge about the permanent properties of objects, since the machinery required to gain a knowledge about certain unchanging properties, such as reflectance and therefore colour, is

¹⁶This handwaving technique, of implying the existence of some functional mechanism, without justification or explanation, is exactly what Zeki criticized about scientists that explained the phenomenon of *colour constancy* (see Section 3.2.4) by ‘discounting the illuminant’, however that might be achieved.

different from the machinery required to gain a knowledge about other unchanging properties, such as form or motion.

3. While it solves one set of problems, functional specialization raises a host of new ones which have to be solved to generate the integrated visual image in the brain. Chief among these is the integration of the results of the operations undertaken by the different specialized areas in order to generate the unified visual image in the brain. The problem of integration is therefore a consequence of functional specialization. It demands that cells in different visual areas, registering different attributes of the visual scene, interact with one another.
4. This raises the problem of what anatomical and functional strategy the visual cortex uses for this interaction. All the anatomical evidence suggests that:
 - (a) there is no single area to which all the specialized visual areas project, and
 - (b) where two areas specialized for different attributes of the visual scene send outputs onto a third area, each maintains its own territory in the third, suggesting that any convergence must be brought about by local circuits.

Physiologically, the problem of integration demands that the cells respond with some kind of temporal synchrony, but this raises the question of who monitors the synchrony, a problem which also imposes itself when one asks who it is who monitors or 'sees' the integrated visual image.

5. One way of solving this problem is for the cells in the specialized visual areas to communicate directly and reciprocally, not only with each other but also with cells in areas V1 and V2 from which they receive their inputs. Such a postulated mechanism receives powerful support from the anatomical evidence which shows that the specialized visual areas are indeed reciprocally connected with one another and send diffuse return outputs back to areas V1 and V2. These reciprocal connections form the anatomical basis for feed-back.
6. Through feed-back, cells could establish groups or repertoires made up of cells in several different visual areas, all of which fire in temporal synchrony. Such groups will become stable if the chances of their responding in synchrony is high and unstable if the chances

are low, with many gradations in between the two extremes. There is a sufficient degree of plasticity in the brain, even in the adult brain, to allow this to occur. It follows from this that some groups may be transient while others are more stable, and that one set of cells may take part in one group at one time and in another group at another time.

7. It follows from this, too, that the synthesis of the visual image in the brain depends not only on the simultaneous activity of cells in the different specialized visual areas, but also on the temporal synchrony of their responses. Moreover, this synchronous activity must also include areas such as V1 and V2 with which the specialized visual areas are reciprocally connected. From which it follows that the visual percept does not reside in any given visual area, even if that area is critical for certain features of that visual image. Rather it is the result of on-going activity in several reciprocally connected visual areas.
8. A review of the experimental and clinical evidence suggests that, for the conscious perception of a visual stimulus and thus for the acquisition of knowledge about the visual world, the simultaneous activity of many visual areas is necessary and that a stimulus will not reach visual awareness unless this condition is satisfied, even if signals reach the specialized visual areas indirectly, by by-passing area V1. This is what is observed in blindsight, a condition in which subjects have no conscious awareness of the stimuli presented to their 'blind' fields and consequently have no knowledge of these stimuli, even if they can discriminate them. One can postulate therefore that blindsight is the consequence of a condition in which there is no synchronous firing of cells in the specialized visual areas with their counterparts in area V1, because the latter is damaged. More simply, the groups or repertoires necessary for the conscious perception of vision cannot be formed.

8 Conclusion

I believe that studying Zeki's "A Vision of the Brain" gave me a well-rounded understanding of many aspect of visual perception. I also believe, and have made mention of this in several places, that Zeki's theory is in many places compatible with, or even supportive of, other theories of vision. It is my personal opinion, that hard, physical evidence is worth more, or at least as much, as the most well contemplated thought-experiment. For this reason, it is comforting to know where other theories of vision, mostly based on behavioral studies, can be correlated to

corresponding physiological studies. For example, while reading Hoffman [3], I would always wonder if there is other evidence to back his claims, or if there could be other explanations, which are just as permissible. While I do not take the entirety of Zeki's reasoning for granted (he sometimes seems to go off on a tangent), he does by no means exclusively propose his own findings. Scores of references link Zeki's own work to that of others, proving that the arguments he offers or cites are used and accepted throughout a large portion of the scientific community.

Apart from answering many questions about visual perception, Zeki's book also inspired many new questions in me. This is all the more important, as these are intricately related to the research goal of my Ph.D. In this context, I hope that the knowledge gained from Zeki will help me understand what needs, and can be done, in the field of visual perception as it relates to computer graphics.

References

- [1] FREEMAN, W. T., ADELSON, E. H., AND HEEGER, D. J. Motion without movement. *ACM Siggraph 25*, 4 (August 1991), 27–30.
- [2] GORDON, I. E. *Theories of Visual Perception*. John Wiley & Sons Ltd., 1989.
- [3] HOFFMAN, D. D. *Visual Intelligence: How We Create What We See*. W.W. Norton & Company, NY, 2000.
- [4] LAND, E. H. The retinex theory of color vision. *Proceedings of the Royal Institute of Great Britain 47* (1974), 23–58.
- [5] ZEKI, S. *A Vision of the Brain*. Blackwell Scientific Publications, Oxford, 1993.

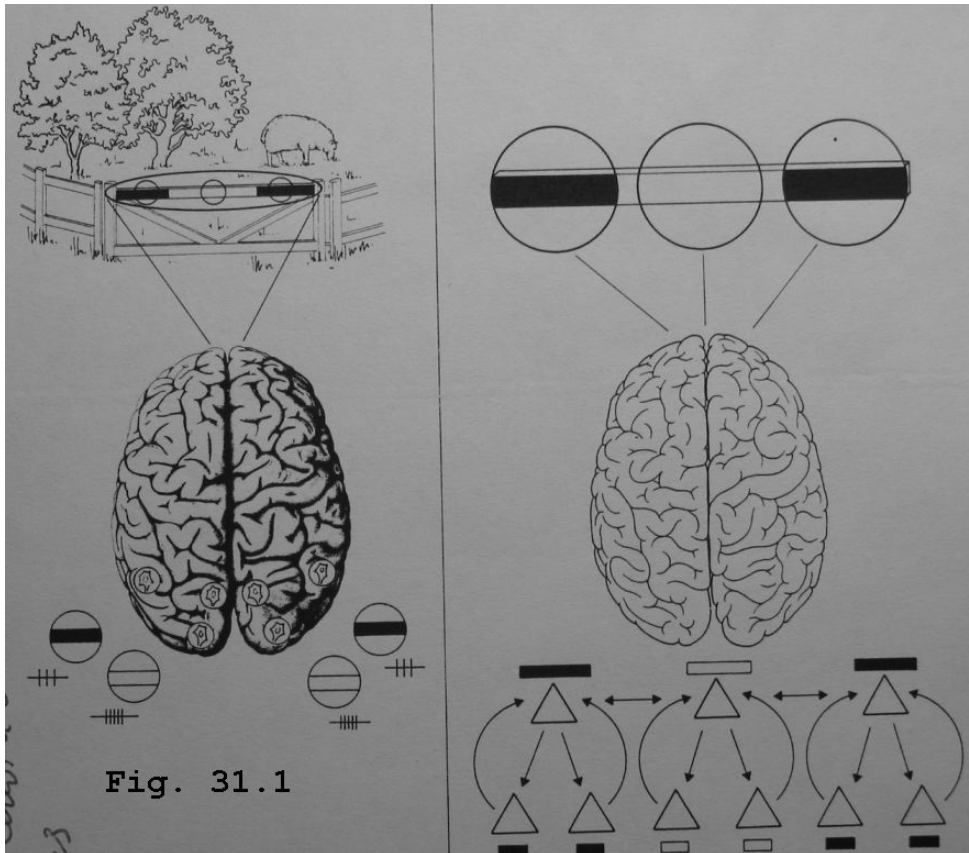


Figure 8: Original Caption: “Some mechanism must ascertain that cells in the cortex responding to the same orientation but with different receptive fields are responding to the same object. This figure shows one possible system of connections to ensure this. The cells of V1 (bottom row, right) registering the horizontal orientation but with non-overlapping receptive fields communicate with one another, and also with cells in V3 (top row, right) registering the same orientation. The latter not only communicate iwth each other but also have a return input back onto the cells of V1 which feed them.” (p.322)

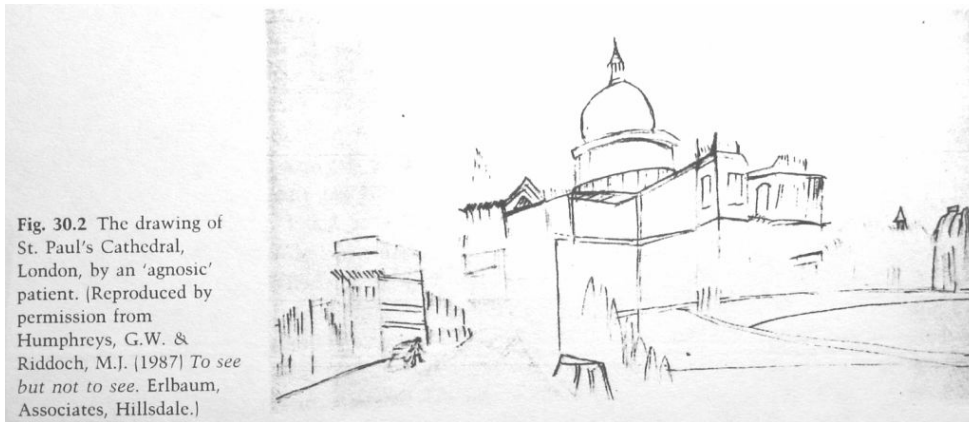


Figure 9: see original caption text

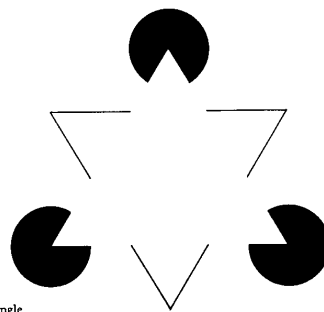


Fig. 30.3 The Kanizsa triangle.

Figure 10: Most people see in this figure a white triangle that overlays an arrangement of three circles and another, outlined triangle. As the white triangle is only suggested or induced by the particular arrangement and form of the other shapes, it is sometimes called an *illusory figure*. Agnosic patients can often not perceive illusory contours or figures.