Visual Information Processing and Phenomenal Consciousness*

(In: T. Metzinger (Hg.) *Conscious Experience*.
Paderborn: Schöningh/Imprint Academic 1995, 409-424)

Ansgar Beckermann

I

As far as an adequate understanding of phenomenal consciousness is concerned, representationalist theories of mind which are modelled on the information processing paradigm, are, as much as corresponding neurobiological or functionalist theories, confronted with a series of arguments based on inverted or absent qualia considerations. These considerations display the following pattern: assuming we had complete knowledge about the neural and functional states which subserve the occurrence of phenomenal consciousness, would it not still be *conceivable* that these neural states (or states with the same causal rôle or the same representational function) occur without having a phenomenal content at all, or that these states are accompanied by phenomenal contents differing widely from the usual ones.

I think these arguments are *prima facie* fairly plausible, but in the case of representationalist theories they are plausible only if one confines oneself to the representational states themselves and completely ignores the specific way in which the representations in question come about. I would like to exemplify what I mean by using the case of visual perception.

Let us assume, Harvey sees that a glass is on the table in front of him. How would a representationalist analyse this state? Well, although we normally find analyses of types of mental states like beliefs and desires in representationalist theories, for all I know they generally do not concern themselves with perceptual processes. However, it does not seem implausible to assume that for a representationalist perceptual processes are mainly processes of acquiring beliefs – hence (s)he will mainly focus on the beliefs in which the perceptual processes result. And for this type of mental state (s)he has an analysis. If Harvey sees that a glass is on the table in front of him, then this state consists, amongst other things, in Harvey's standing in a certain functional/computational relation *R* to a mental representation *mr* with the content that there is a glass on the table in front of Harvey – or, put casually, that the mental

---

* I would like to thank Achim Stephan for his helpful comments on an earlier draft of this paper and Antonia Barke for translating the paper into English.

representation *mr* is in Harvey's belief box. In order to distinguish perceptual beliefs from other beliefs however, one would have to say a little more about how *mr* gets into Harvey's belief box – e.g. that *mr* is more or less directly caused by the very fact, which features as its content, and moreover that in this causal process Harvey's eyes and the light reflected by the objects involved play a crucial rôle. A representationalist analysis of the state that Harvey sees a glass on the table in front of him could, therefore, look like this:

(a) In Harvey's belief box there is a mental representation *mr* with the content that on the table directly in front of Harvey there is a glass.

(b) That this mental representation is now in Harvey's belief box is, amongst other things, caused by the fact which features as its content, and in this causal process Harvey's eyes and the light reflected by the objects involved play a crucial rôle.

However, if one takes this analysis as a starting point, one immediately encounters the problems mentioned above, because if the state that Harvey sees a glass on the table in front of him is exhausted by (a) and (b), it seems perfectly conceivable that this state does not have any phenomenal content or that it is accompanied by a phenomenal content differing widely from the usual one. However, in my opinion this is only due to the fact that in point (b) *far too little* is said about the manner in which Harvey's perceptual belief comes about. That is to say, the phenomenal content of perceptual states can only be explained adequately within the framework of a representationalist theory of mind if one does not focus exclusively on the mental representations, but also analyses very carefully the manner in which these representations come into being. In order to illustrate this thesis, I shall outline the essential aspects of visual information processing as it is at present seen by the cognitive sciences and AI research.

II

Visual information processing begins with retinal images (or more precisely: distributions of electrical impulses of the receptor cells in the retina), or in the case of artificial systems with raw images generated by a television camera and encoded in two dimensional arrays of pixels. Vision, however, is a sense of distance and has to inform us not so much about proximal, but about distal stimuli. Retinal images or raw images are therefore only interesting if information can be extracted from them about the three-dimensional physical scenes which have caused them. The processing must, therefore, end with descriptions or representations of these scenes. Or, in a shorthand version: visual information processing begins with a retinal image or a raw image and ends with a representation of the objects which make up the corresponding visual scene – the types of these objects, their positions in space, their sizes, their shapes, their surface textures, their motions. This enables Ballard and Brown (1982) to speak with regard to visual information processing of "the construction of explicit, meaningful descriptions of objects from images".

With regard to vision in general, therefore, three main components (Cf. Neumann 1993: 566) have to be distinguished:

- physical objects in a scene
- images of the scene as input of the visual system
- a description or representation of the scene as its output.

These components combine in the following way: in the first step, the physical objects which make up a three-dimensional scene cause an image in the input medium of the system in question, which is neurally or electronically encoded. (Images, in the sense in which this term is understood here, therefore are nothing else than two-dimensional projections of three-dimensional scenes.) In the second step, the task of visual information processing consists in the reversion of the first process, i.e. the reconstruction of the original scene from the image.[1] Visual information processing has to lead to an output that tells us 'what is where' in the scene.

> The 'where' in this expression refers to spatio-temporal information, i.e. the reconstruction of the scene's geometry, while the 'what' implies an interpretation of the content of the scene, and in particular the recognition of objects. (Neumann 1993: 566-7).

So visual information processing consists in the reconstruction of a scene from an image caused by that scene. Or to put it more precisely: It consists in the construction of a representation of a scene from an encoding of an image caused by that scene.

Today it is generally assumed that four stages can be distinguished within the process of reconstruction which constitutes visual information processing, and these four stages are usually grouped into two categories – low-level and high-level vision.[2] Each of these four stages – at least according to Marr (1982) – leads to new

1 However, generally a single image is not sufficent for this reconstruction. In practice one therefore often does not work with individual images, but sequences of images. It is obvious that reconstructions of temporally extended dynamic scenes are only possible on the basis of such sequences.

2 For the following cf. e.g. Neumann 1993: 569-570.

representations, until in the end a representation of the corresponding three-dimensional physical scene results.[3]

The first step in low-level visual processing – leading to a *primal sketch* – aims at *segmentation of the image*. For this, the isolation and representation of the central elements of the picture, such as edges, homogenous areas, texture etc. is crucial.

The second step in low-level visual processing – leading to what Marr calls a *2½D-sketch* – is already concerned with the interpretation of elements of the image as elements of the scene; that is to say, elements of images are correlated with parts of real three-dimensional scenes. For example: An edge in the image is interpreted as a shadow border, a red area as a wall of a house, a green textured area as a grassy patch etc.

The first step in high-level visual processing concerns the *recognition of objects*. On the basis of the segmented representation of the two- and three-dimensional structure of the image made available by low-level vision the high-level processes must first determine what objects are present in the scene, which properties these objects have and in which ways they are interrelated. In general this presupposes substantial knowledge of what objects look like when seen from different perspectives, because recognition of objects is essentially the reversal of the image-creating process.

Further stages of high-level visual processing are sometimes summarized under the catchphrase 'high-level image interpretation'. "They generally aim at recognizing relationships across object and time boundaries, e.g. interesting configurations of objects, special situations, cohering sequences of movements etc. Similar to the process of object-recognition model-like knowledge of what one wants to recognize plays an important rôle." (Neumann 1993: 567)

Schematically then, the individual stages of visual information processing can be summarized as follows:

---

[3] The following four stages are at least taken as a basis by AI scientists in the construction of artificial visual systems. Whether they can be found in the same way in natural systems is not entirely clear. Regarding this point also cf. section III below.

raw image

↓

image segmentation

↓

low-level interpretation of picture elements

↓

object recognition

↓

high-level image interpretation

III

In order to economize the use of storage capacity, it is an aim of AI research to organize the above mentioned stages in such a way, that later steps of image processing can fulfill their task, as far as possible, on the basis of the results of low-level visual processing without recourse to raw images. A 10 second sequence of color TV pictures has the considerable data volume of 220 GByte. Hence, it appears sensible to burden the available storage as little as possible with raw images.

However, it seems to be characteristic of *human* visual information processing that raw images are not simply discarded after the first steps of processing. If we can rely on introspection, it even seems that the processing of retinal images does not generate successively new *independent* representations, but rather that the results of the individual processing steps are closely interrelated and also each leads to an improvement or better focussing of the original image. Once the first step of low-level processing has been completed, the homogenous areas appear more clearly, the borders between these areas are more distinct; the image itself looks 'sharper'. With the second step of low-level vision the interpretation of the image begins: we cease to see white, grey or differently colored areas or differently shaped edges, but instead the surface of a table, a shadow or a green-grey structured background. It is as if labels have been stuck on the elements of the image indicating which elements of the scene they are correlated with. After the stage of object recognition, we no longer have the impression of seeing an image; instead we look – so to speak – through the image onto the objects, which caused it.[4] Now we see a table, a glass, patterned wall

---

4 In this context Van Gulick (cf. e.g. 1989: 223ff.) speaks of the high 'semantic transparency' of phenomenal representations. Although this expression certainly is highly suggestive, it does

paper etc. This step, too, usually leads to a clearer focussing of the original image, because the knowledge of how the perceived objects normally look under the given conditions, obviously is used to complete and improve the original image. Finally, we even see the properties and spatial relations of the objects perceived, e.g. that the table is white, that the glass is on the table, etc. So, introspection seems to bear witness to that there are indeed stages in our perceptual process, which correspond to the four stages of visual information processing which have been postulated by cognitive scientists.[5]

However, these stages do not lead to representations which are independent of each other. This is not to say that not every one of these stages leads to new representations, but simply that these representations are not independent – rather, every one of them also modifies the preceding representations and all of them lead to an alteration and better focussing of the original image.

As already mentioned, in the end we even have the impression of seeing objects and scenes directly, instead of just seeing an image. And this impression to a certain extent is surely right – to a certain extent, however, the image still seems to be there too. It is to this fact that sense data theorists have tried to draw our attention time and again. Moore, for example, invites us to observe carefully what happens when we press one of our eyeballs with one finger while not changing the perspective which we have onto the scene. Basically, there is no change in the perceived scene – we still see the same objects in the same configuration. However, something changes after all: the original image of the scene. Or, to put it more precisely: the elements of the image (homogenous areas and edges) which are caused by the individual objects, change their shapes. The fact that Moore points to, is something which all of us are familiar with: we normally 'see through the images' the objects and scenes, which have caused these images, but we *can* also concentrate on the images and their

---

not correspond entirely with the ideas put forward in this paper. For, according to Van Gulick, the notion of semantic transparency concerns the "extent to which a system can be said to understand the content of internal symbols or representations on which it operates" (223) However, this presupposes that perceptual images are representations and this presupposition in my opinion is far from being obvious (cf. section IV, below). Moreover, Van Gulick comes dangerously close to a homunculus-fallacy - for *who* is supposed to understand the meanings of its internal representations? It is an advantage of the account proposed in this paper that it entirely avoids such homunculus-presuppositions, because according to this account the 'semantic transparency' of perceptual images simply consists in the fact that the system closely links these images to the explicit representations of the perceived scenes.

5 The fact that the process of visual information processing in humans also passes through different stages, however, is shown clearly only in certain psychological experiments, in which perceptual stimuli are shown for a few milliseconds.

elements. In other words: at the end of the perceptual process the original images are still there – albeit in a modified form.

IV

Looking back over the argumentation so far, the reader will hardly be surprised that my central thesis is as follows: if visual information processing is structured in the way I have just described, i.e. if during this processing representations are reconstructed from the original images of a scene in such a way that in the end these images are not discarded, but only modified, then this kind of processing has a phenomenal aspect. This is at least true if the modified original images are as accessible to the system in question as the explicit representations of the scene perceived. Thus, it is crucial for the phenomenal character of these perceptual processes, that in addition to the explicit representations of the perceived scene, the original images are preserved.

With regard to this thesis, one can of course again raise the question whether it is not at least conceivable that in a system visual information processing of the kind described takes place without having a phenomenal content at all, or that these processes are accompanied by phenomenal contents differing widely from the usual ones. Before trying to give a direct answer to this question, however, I would like to remark on three issues which, as I hope, will serve indirectly to increase the plausibility of my central claim. The first of these remarks concerns the relation of my thesis to positions which can be found in the literature on this subject.

In the literature, the opinion is widespread that within the framework of a representationalist theory of mind, the existence of phenomenal states can and must be explained through certain types of representations – viz. analog or pictorial representations.[6] The connection between this position and my claim is obvious. At least it is obvious if one does not make a difference between pictures and analog representations. As regards the latter question, however, I would like to defend an opinion which is in exact opposition to the majority view.

While most authors treat pictures as paradigm cases of representations, images of the kind in question here are nothing more than two-dimensional projections of three-dimensional scenes. That is to say, there is no *prima facie* reason to regard them as representations at all. *A fortiori*, pictures are not a special kind of analog representations. Indeed, pictures and analog representations have much less in common than is usually thought. Though not being representations, pictures can be

---

[6] Cf. e.g. Nelkin 1989, 1994.

'interpreted' however, because of the regular connections between them and the scenes whose projections they are. That is to say, scenes can be reconstructed from pictures. Moreover, for our visual system this reconstruction is an especially natural task, since external pictures – *cum grano salis* – generate the same retinal images as the depicted scenes. Since it is the main task of our visual system to interpret retinal images (cf. above), the interpretation of external pictures does not normally present any problem.

However, the difference between pictures and (analog) representations is not crucial in this context, because the majority of authors who wish to explain phenomenal states within the framework of a representationalist theory of mind through a certain type of representation, have in mind image-like or pictorial representations. So there seems to exist a common intuition that images or pictorial representations may well play a crucial rôle in the explanation of phenomenal states, and especially so, if the explanation of visual impressions is at issue.

Another position to which I have a great affinity has been developed by N. Humphrey in his recent book *A History of the Mind*. According to this position, in their development of the ability to perceive animals have, in the course of evolution, developed two distinctly different representational systems.

Since animals exist, the outside world impacts their boundaries (membranes, skins). Light falls on them, objects bump into them, pressure waves press against them, chemicals stick to them. Some of these events are 'good things' for the animal, others are bad. Hence, it constitutes an evolutionary advantage if the animal has the means to sort out the good events from the bad ones, or, in order not to complicate matters too early, to react differently in different cases. 'Natural selection was therefore likely to select for "sensitivity".' (Humphrey 1993: 18).

At first these reactions are still local: the surface contracts or secretes certain chemicals. An important new stage is reached when signals are transmitted from one part of the surface to other parts so that reactions also may take place there. For this enables responses that are better adapted to the animal's needs: for example it may swim away, rather than just recoil from a noxious stimulus.

But even at this stage there is still a fairly direct connection between stimulus ('perception') on the one hand and reaction ('action') on the other. The next stage is reached when sensitivity and responsivity become partially decoupled. Signals from parts of the surface will be transmitted, but they no longer result automatically in definite reactions. Whether, and in what way, the animal reacts, rather depends on a number of other factors, which – one could say – are taken into account in a central process of determination.

At this stage it is not yet necessary to store the signals transmitted from the surface or to represent them permanently. However, this becomes indispensable if it proves necessary to detach the reaction *in time* from the stimulus. Thus, a first system of representations is developed, if not only the sensory side and the response side of the process are partially decoupled, but there also is a delay between stimulation and response. A central site has to evolve where representations of stimuli can be stored for a certain amount of time.

One answer to the question of what these representations represent, seems to suggest itself: the proximal stimuli which disturb the boundaries of the animal – the light which strikes the surface, the objects which bump into the animal, the chemical substances with which it comes into contact. However, if we examine the matter a little more closely, it becomes clear that the information transmitted from the surface is less dependent upon the proximal stimuli and more upon the effect of these stimuli on the surface. If two different objects affect the surface in the same way, the same information will be transmitted; if the same object has a different effect at different times, the information transmitted will be different. In other words: the first central representations, which in the course of time take shape in animals, carry information about the state of the animal itself, rather than about proximal stimuli, or, to put it more philosophically: they carry information about how it is for the animal when its surface is affected by certain proximal stimuli. It is, therefore, not surprising that Humphrey identifies these representations with 'raw sensations'.

> So the phenomenology of sensory experiences came first. Before there were any kinds of phenomena there were 'raw sensations' – tastes, smells, tickles, pains, sensations of warmth, of light, of sound and so on. (Humphrey 1993: 21)

Thus the first representations only provide animals with answers to the question 'what is happening to me?'. However, for the survival of most animals it is undoubtedly just as important to know what is happening around them. Hence, answers to the question 'what is happening out there?' are of the utmost interest. How can representational systems be developed which provide answers to this question?

In a nutshell, Humphrey's claim is that in the course of evolution, in addition to the first representational system, a second one developed, which also takes the signals transmitted from the surface as a point of departure, but processes these signals in a radically different way.

> By the end of the first stage of evolution sense organs existed with connections to a central processor, and most of the requisite information about potential signs was being received as 'input'. But the subsequent processing of this information, leading to subjective sensory states, had to do with quality rather than quantity, the transient present rather than permanent identity, me-ness rather than otherness. In order that the same information could now be used to represent the

> outside world, a whole new style of processing had to evolve, with an emphasis less on the subjective present and more on object permanence, less on immediate responsiveness and more on future possibilities, less on what it is like for me and more on how what 'it' signifies fits into the larger picture of a stable external world.
>
> To cut a long story short, there developed in consequence two distinct kinds of mental representation, involving very different styles of information processing. While one path led to the qualia of subjective feelings and first-person knowledge of the self, the other led to the intentional objects of cognition and objective knowledge of the external physical world. (Humphrey 1993: 22)

The main point of Humphrey's theory can therefore be summarized as follows:

(a) In the course of evolution, two fundamentally different ways of processing stimuli which affect the surface of animals developed ('sensation' and 'perception'). The first leads to representations, which encode answers to the question 'what is happening to me?', the second to representations which answer the question 'what is happening out there?'.

(b) Representations of the first kind account for the phenomenal aspect of perceptual processes.

(c) Although both types of information processing are usually closely linked in the various perceptual processes, they nevertheless remain different and can, therefore, occur independently of each other.

Without any doubt there are certain differences between the model developed in this paper and Humphrey's theory – however, there are obviously many parallels between the accounts, too. E.g. representations of the original images, which according to the position developed in this paper are responsible for the phenomenal aspect of visual perception, can without any problem be seen as representations of what happens on the retina – i.e. as representations of the first type of Humphrey's theory. Since the fundamental affinity of the two accounts is clear already from these sketchy remarks, I shall not pursue these details here.

The second remark concerns the much discussed question of which evolutionary advantage phenomenal consciousness might offer. In this paper, of course, I cannot answer this question in a general manner, but I would like to mention a very important aspect: the rôle of perceptual images in the control of behavior, which can be demonstrated even through very simple examples. Everyone knows the situation in which the task is (e.g. on a motorway) to maintain a constant distance with the car in front. This can be achieved by calculating, as part of the process of object recognition, i.e. *during high-level visual processing*, the distance between one's own car and the car in front and by accordingly increasing, lowering or maintaining a constant speed, depending upon whether this distance has become bigger, smaller or remained the same. However, the same effect can also be achieved with much less effort if only the system tries to ensure that the *size* of the *picture* of the car in front stays unchanged, i.e. if it increases, lowers or maintains a constant speed, depending upon whether the size of the picture shrinks, grows or remains unchanged. That is to say, the same behavior which is adapted to the situation can be generated without recourse to the results of high-level visual processing, but only by utilizing the results of the first step of low-level visusal processing, i.e. with reference to the elements of the image. Moreover, it seems plausible to assume that the second way is much quicker and generally more reliable. Even finding one's way through a room with chairs and tables and avoiding obstacles in general can often be regulated with the help of a relatively simple image analysis.

As a second example let us take the action of grasping something. One can obviously control this action, too, by using the results of high-level visual processing. First one determines the coordinates of the object one wants to grasp and then designs a plan with the help of which muscle contractions one can move one's hand to these coordinates. However, this is a lengthy and – e.g. if one attempts to grasp a moving object – very complex process often with an unreliable outcome. How difficult this type of behavioral control really is, can be illustrated by imagining that one has to grasp an object with *closed eyes*, only on the basis of information about the location of the object and one's hand and which effect certain muscle contractions have on the position of the latter.

Obviously, human behavioral control works in a completely different way. How central a rôle perceptual images play in this control becomes apparent through the fact of how badly this control functions when one closes the eyes or is blindfolded. Therefore, it seems not too far fetched to claim that the grasping of objects as well is largely controlled through perceptual images and their elements. The details are not as yet entirely clear, but one can easily imagine that in grasping we do not aim for congruency of the coordinates of our hand and the object we want to grasp, but rather at bringing the *image elements* which are produced by my hand and the object closer. (Additionally there would obviously have to be a mechanism which takes account of the dimension of depth.) Maybe the image element of the object even acts as a kind of attractor which 'attracts' the image element of my hand.[7] On the whole, it seems very likely to me that in planning actions, we do not only use our knowledge of which bodily movements have what effect *in the world*, but we especially use our knowledge of how *our perceptual images* change as a result of these movements.

---

[7] This interesting idea I first encountered in a discussion with the psychologist Michael Stadler (Bremen).

The efficiency of bodily movements can often already be judged on the basis of the analysis of these images – as shown in the above mentioned example where the task consisted in maintaining a constant distance with the car in front.

Finally, the third remark concerns the fact that the position put forward here, accords well with familiar neurobiological findings. As is well known, the axons of the retinal ganglion cells are collected into the optic nerves, which leave the eyes roughly at the height of the fovea. The optic nerves converge at the base of the skull and exchange half of their fibres at the optic chiasma. After this re-sorting process the optic nerves travel to the lateral geniculate nuclei. These nuclei contain approximately the same number of neural fibres as the optic nerves, and these fibres project directly onto the primary visual cortex in the occipital lobe (area 17). Area 17 in turn projects in an ordered manner, i.e. in the form of a one-to-one mapping, onto the secondary visual area (area 18), which itself projects onto at least three more structures: onto a region called MT, onto the tertiary visual area (area 19) and onto another visual area called V4. This kind of structuring continues – one area always projects onto a few others. In addition, every single one of these areas passes back signals to the areas from which it receives inputs. Moreover, the individual areas additionally project onto deeper structures of the brain – e.g. the superior colliculi and different parts of the thalamus. Finally, all the visual fields receive inputs from sub-structures of the thalamus: just as the lateral geniculate nuclei project onto the primary visual cortex, other parts of the thalamus are connected to other areas.

In this context two points are of particular relevance. Firstly, the fact that the projection of the axons of the retinal ganglion cells – even after switching over in the lateral geniculate nuclei – are organized in a *retinotopic* manner. Adjacent ganglion cells project onto adjacent parts of area 17, so that the topological structure of the firing patterns of the ganglion cells remain intact under this projection. In other words, *the firing pattern of the neurons in the primary visual cortex, too, can be taken as an encoding of the retinal image*. Secondly, the fact that later visual areas send signals back to the areas from which they receive inputs, can easily be linked to the observation that later stages of visual information processing, amongst other things, always lead to an alteration and improvement of the original image. In my view, it therefore does not seem implausible to assume that the primary visual cortex is the physical correlate of the phenomenologically relevant perceptual image (or at least constitutes an essential part of this physical correlate).

This assumtion is also supported by pathophysiological findings. It has been known for a long time that small lesions, limited infarcts or small tumors in the primary visual cortex lead to blindness of a circumscribed part of the visual field, or, if the whole of area 17 is affected, to complete blindness. One refers to these cases as cortical blindness in order to distinguish them from instances in which the reason for the blindness is a deficiency in the eyes or the optic nerves. Obviously, cortical blindness is exactly what is to be expected if the primary visual cortex really is the physical correlate of the phenomenal perceptual image. Any damage of the correlate ought to lead to a loss within this image and to a failure of all higher representations which depend upon a proper functioning of this perceptual image.

In the last few years the phenomenon of 'blindsight' has been much discussed in this context. This phenomenon manifests itself particularly in the fact that patients suffering from cortical blindness and therefore say that they literally do not see anything in some areas of their visual field, nevertheless possess certain 'information' about what they claim they do not see. They e.g. give under forced choice conditions many more correct answers than could statistically be expected if there were a complete lack of information. In the model outlined above this effect could be explained as follows: later processing stages which, due to neural wiring, receive inputs not only from the primary visual cortex (or later areas), but also from the thalamus, can still fulfil their respective tasks (at least partially). However, in these cases there is no phenomenal perceptual image since this image is tied to the primary visual cortex which is seriously damaged. And it is this fact that accounts for the effect that the person in question – subjectively completely justified – reports not having seen anything.

Other phenomena concerning split-brain patients whose corpus callosum has been cut can also be explained within the framework of this model. If an object is presented only in the left half of these patients' visual field, they cannot verbally answer the question of which object is shown to them. This is due to the fact – as is generally assumed – that stimuli of the right half of the retina are only transmitted to the right hemisphere and that in the case of split-brain patients, therefore, the parts of the brain involved in the production of language in the left hemisphere receive no (direct) information about the objects which are only shown in the left half of the visual field. Within the framework of the model suggested here, this would mean that split-brain patients to a certain extent possess an intact perceptual image of the objects presented, but that higher (especially language related structures) have no access to this image and that therefore the patients lack the ability to give adequate answers.

V

Let us return to the crucial issue of whether it is plausible to think that processes of visual perception which are structured in the way described above, always have a phenomenal aspect. Why should the fact that perceptual images play a central rôle in these processes make all the difference?

If two systems *A* and *B* are functionally equivalent, i.e. if they have the same functional structure, and, at time *t*, are in identical functional states, then this implies that the two systems display exactly the same behavior at *t*. This is so, because functional states are characterized by their causal relations to inputs, outputs and other functional states. If two functionally equivalent systems are in identical functional states, the same inputs therefore cause identical internal processes, which in the end must lead to the same behavioral output. Proponents of 'absent' or 'inverted qualia' arguments must, therefore, accept the assumption that it is possible that two systems whose behavior is indistinguishable, nevertheless differ radically with regard to the phenomenal content of their respective states. Or (yet stronger) that it is even possible that of two systems which behave exactly alike, only one possesses phenomenal states at all, while the other lacks these states completely.

Despite all the difficulties it implies, this consequence has been accepted by many philosophers, since it seems intuitively plausible to claim that behavioral evidence never suffices to decide in which phenomenal state a given system is. However, as early as 1975 S. Shoemaker pointed out that the situation begins to look very different if two systems do not just display the same behavior, but also agree in all *beliefs about their own states* in general and their own phenomenal states in particular. Proponents of the view that even the totality of behavioral evidence is not sufficient to judge which phenomenal content certain states have, usually hold at the same time that another, direct way exists to decide this question – *introspection*. However, according to Shoemaker introspection depends upon what a being believes or knows about its own states. Hence, if two beings do not just both behave in a way characteristic for beings which are in pain, but also both believe in the same manner that they feel pain, then not only all the behavioral evidence supports the point of view that they really feel pain, but also all introspective evidence. And what other evidence could there be for the claim that at least one of these beings does not feel pain?

> So one way of putting our question is to ask whether anything could be evidence (for anyone) that someone was not in pain, given that it follows from the states he is in ... that the totality of possible behavioral evidence *plus* the totality of possible introspective evidence points unambiguously to the conclusion that he is in pain? I do not see how anything could be. (Shoemaker 1975: 189-90)

If two systems not only correspond in their behavior, but also in all beliefs concerning their own states in general and their own phenomenal states in particular, it is impossible – at least from an epistemic point of view – to differentiate between them as regards the phenomenal content of their mental states.

However, a system whose visual information processing is organized in the way described above, and which additionally possesses meta-representations concerning not only the representations of the perceived scenes which form the final result of this processing, but also concerning its earlier stages and especially representations of the modified original image, exactly fulfills this condition – at least with respect to the phenomenal aspects of *visual* perception. In any case when questioned the system will give answers not only about the objects and situations which are accessible to it on the basis of its visual apparatus, but also about the visual impressions it has gained in this context. It will not only say that the table in front of it is square and that a glass is on the table; but e.g. also say that the upper rim of the glass looks elliptical although it is round, that the surface of the table looks slightly reddish or that the wallpaper in the background looks so faded that it cannot discern its pattern.

If one attributes beliefs to this system (and I cannot conceive of any reason not to do this), these beliefs do not just refer to its environment, but also to its own states and in particular to something which *we* would call visual impressions and which the system itself describes *in exactly the way* in which we describe our own visual impressions. So what kind of evidence could there be from which we could conclude that the system, contrary to ourselves, does not possess visual impressions? Evidence that would not at the same time lead us to doubt that our fellow-humans, or indeed we ourselves, have visual impressions?

This point is closely connected to a second one. Systems like the one described can distinguish between how things are, and how they seem to be, in exactly the same way as we can. Some philosophers, inspired by Wittgenstein's later thought, hold that sentences like 'the wall looks red' as well as 'the wall seems to be red' have nothing to do with any phenomenal qualities; rather these sentences ought to be understood in a performative sense. According to those philosophers, when I utter these sentences I only indicate that I am not prepared to commit myself to the truth of my statements. If anyone believes what I say on the basis of my utterances, he or she does so at his/her own risk. By contrast when I say 'the wall is red' or even 'I know that the wall is red', I am accountable for the content of these statements, and I therefore run the risk of being held responsible if things turn out to be different from how I say they are.

R. Chisholm, however, has shown that the performative use of 'to seem' and 'to look' at best constitutes one possible use (cf. Chisholm 1989: 20-22). There are also other uses which after all seem to be related to phenomenal qualities; there is e.g. a way of speaking in which it makes sense to say:

(1) The wall seems to me grey in this light, but I know that it is red.

And even:

(2) The wall seems to me red in this light, and I know that it is really red.

Especially in the second example 'to seem' cannot have the performative use as mentioned above, because if this were the case, the second part of the sentence would defeat the point of the first. Obviously this is not the case. However, if 'to look' or 'to seem' in these examples are not used in the their performative sense – in what sense then *are* they used?

In order to answer this question, it is important to ascertain in which situations we use 'to seem' or 'to look' in the non-performative sense of example (2). The paradigmatic case in my opinion is this: in a psychophysical experiment I am requested by the experimenter to attend exclusively to the color which a wall in front of me *seems* to have in different lighting; a wall which I could satisfy myself earlier to be white. What does the experimenter want to know from me? Obviously he is not interested in my beliefs as to the real color of the wall, for I know that the wall is white and this belief does not change in the course of the experiment. Hence, he wants to learn about the subjective *impressions* which I have under different experimental conditions. If he asks me which color the wall *seems* to have, he assumes that not my objective beliefs, but my visual impressions will vary during the experiment. It is these changes that he wishes to know about.

This brings us back to the point I have already mentioned at the end of section III. Due to the manner in which our perceptual system is organized, we are not only able to say something about the perceived objects and scenes, but we can also concentrate on the (modified) original image and report how this image changes without thereby implying that these changes presuppose changes in the perceived scene itself. However, this also means that every system in which the process of visual information processing is structured in the way outlined above, can also distinguish between changes regarding the perceived scene and changes which only concern the *image* of that scene, i.e. how this scene *appears* to the system. In this sense the distinction between how things are, and how they just seem to be, is entirely natural for a system like that.

To sum up: The thesis put forward in this paper is that perceptual processes in which representations of real three-dimensional scenes are constructed from raw images in such a way that in the end these images are not discarded, but only modified, always have a phenomenal aspect. And the main argument for this thesis is that systems whose visual information processing is organized in this way, do not only behave in exactly the same way as we do (e.g. grasp objects in the same quick and elegant manner), but also have the same beliefs with regard to themselves and their perceptions as we do, or at least talk about themselves and their perceptions in the same way as we do. Just like ourselves, they can distinguish between perceived scenes and visual impressions, just like ourselves they can differentiate between how things really are and how they only seem to be. What could, therefore, count in favor of the claim that these systems nevertheless do not possess mental states with phenomenal content?

References

Ballard, D. and Brown, C. (1982) *Computer Vision*. Englewood Cliffs, NJ: Prentice Hall.

Chisholm, R. (1989) *Theory of Knowledge*. 3rd ed. Englewood Cliffs, NJ: Prentice Hall.

Humphrey, N. (1993) *A History of the Mind*. London: Random House.

Marr, D. (1982) *Vision*. San Francisco: Freemann & Co.

Nelkin, N. (1989) Unconscious Sensations. *Philosophical Psychology* 2, 129-141.

Nelkin, N. (1994) Phenomena and Representation. *British Journal for the Philosophy of Science* 45, 527-547.

Neumann, B. (1993) Bildverstehen – ein Überblick. In: G. Görz (ed.) *Einführung in die künstliche Intelligenz*. Addison-Wesley: Bonn/Paris/Reading, MA, 559-588.

Shoemaker, S. (1975) Functionalism and Qualia. *Philosophical Studies* 27, 291-315. Reprinted in S. Shoemaker, *Identity, Cause, and Mind*. Cambridge: Cambridge University Press 1984, 184-205.

Van Gulick (1989) What Difference Does Consciousness Make? *Philosophical Topics* 17, 211-230.