BEHAVIORAL AND BRAIN SCIENCES (2001) **24,** 939–1031
Printed in the United States of America

A sensorimotor account of vision and visual consciousness

J. Kevin O'Regan
Laboratoire de Psychologie Expérimentale, Centre National de Recherche Scientifique, Université René Descartes, 92774 Boulogne Billancourt, France
oregan@ext.jussieu.fr http://nivea.psycho.univ-paris5.fr

Alva Noë
Department of Philosophy, University of California at Santa Cruz, Santa Cruz, CA 95064
anoe@cats.ucsc.edu http://www2.ucsc.edu/people/anoe/

Abstract: Many current neurophysiological, psychophysical, and psychological approaches to vision rest on the idea that when we see, the brain produces an internal representation of the world. The activation of this internal representation is assumed to give rise to the experience of seeing. The problem with this kind of approach is that it leaves unexplained how the existence of such a detailed internal representation might produce visual consciousness. An alternative proposal is made here. We propose that seeing is a way of acting. It is a particular way of exploring the environment. Activity in internal representations does not generate the experience of seeing. The outside world serves as its own, external, representation. The experience of seeing occurs when the organism masters what we call the governing laws of sensorimotor contingency. The advantage of this approach is that it provides a natural and principled way of accounting for visual consciousness, and for the differences in the perceived quality of sensory experience in the different sensory modalities. Several lines of empirical evidence are brought forward in support of the theory, in particular: evidence from experiments in sensorimotor adaptation, visual "filling in," visual stability despite eye movements, change blindness, sensory substitution, and color perception.

Keywords: action; change blindness; consciousness; experience; perception; qualia; sensation; sensorimotor

1. Introduction

1.1. The puzzle of visual experience

What is visual experience and where does it occur?

It is generally thought that somewhere in the brain an internal representation of the outside world must be set up which, when it is activated, gives us the experience that we all share of the rich, three-dimensional, colorful world. Cortical maps – those cortical areas where visual information seems to be retinotopically organized – might appear to be good candidates for the locus of perception.

Cortical maps undoubtedly exist, and they contain information about the visual world. But the presence of these maps and the retinotopic nature of their organization cannot *in itself* explain the metric quality of visual phenomenology. Nor can it explain why activation of cortical maps should produce visual experience. Something extra would appear to be needed in order to make excitation in cortical maps provide, in addition, the subjective impression of seeing.

A number of proposals have come forth in recent years to suggest how this might come about. For example, it has been suggested, from work with blindsight patients, that consciousness in vision may derive from a "commentary" system situated somewhere in the fronto-limbic complex (taken to include the prefrontal cortex, insula and claustrum; cf. Weiskrantz 1997, p. 226). Crick and Koch (1990),

Kevin O'Regan moved to Paris in 1975 after studying theoretical physics in England, to work in experimental psychology at the Centre National de Recherche Scientifique. After his Ph.D. on eye movements in reading he showed the existence of an optimal position for the eye to fixate in words. His interest in the problem of the perceived stability of the visual world led him to question established notions of the nature of visual perception, and to recently discover, with collaborators, the phenomenon of "change blindness." His current work involves exploring the empirical consequences of the new approach to vision.

Alva Noë is a philosopher at the University of California, Santa Cruz. He received a Ph.D. in philosophy from Harvard University and a B. Phil. from Oxford University, and he has been a Research Associate of the Center for Cognitive Studies at Tufts University. He has published articles on topics in the philosophy of perception, philosophy of mind, and other areas, including a previous *Behavioral and Brain Sciences* target article on perceptual completion. He is currently at work on a book on the relation between perception and action, and he is a co-editor of *Vision and Mind: Selected Readings in the Philosophy of Perception* (MIT Press, forthcoming).

 0140-525X/01 $12.50

Llinas and Ribary (1993), Singer (1993), and Singer and Gray (1995) suggest that consciousness might be correlated with particular states of the brain involving coherent oscillations in the 40–70 Hz range, which would serve to bind together the percepts pertaining to a particular conscious moment.[1] Penrose (1994) and Hameroff (1994) suggest that the locus of consciousness might be a quantum process in neurons' microtubules. Edelman (1989) holds that reentrant signaling between cortical maps might give rise to consciousness. A variety of other possibilities that might constitute the "neural correlate of consciousness" has been compiled by Chalmers (1996b).

A problem with proposals of this kind is that they do little to elucidate the mystery of visual consciousness (as pointed out by, for example, Chalmers 1996b). For even if one particular mechanism – for example, coherent oscillations in a particular brain area – were proven to correlate perfectly with behavioral measures of consciousness, the problem of consciousness would simply be pushed back into a deeper hiding place: the question would now become, why and how should coherent oscillations ever generate consciousness? After all, coherent oscillations are observed in many other branches of science, where they do not generate consciousness. And even if consciousness is assumed to arise from some new, previously unknown mechanism, such as quantum-gravity processes in tubules, the puzzle still remains as to what exactly it is about tubules that allows them to generate consciousness, when other physical mechanisms do not.

1.2. What are sensory modalities?

In addition to the problem of the *origin* of experience discussed in the preceding paragraphs, there is the problem of *differences* in the felt quality of visual experience. Why is the experience of red more like the experience of pink than it is like that of black? And, more generally, why is seeing red very different from hearing a sound or smelling a smell?

It is tempting to think that seeing red is like seeing pink because the neural stimulation going on when we see something red is similar to that underlying our perception of pink: almost the same ratios of long, medium and short wavelength photoreceptors will be stimulated by red and pink. But note that though this seems reasonable, it does not suffice: there is no *a priori* reason why similar neural processes should generate similar percepts.[2] If neural activity is just an arbitrary code, then an explanation is needed for the particular sensory experience that will be associated with each element of the code. Why, for example, should more intense neural activity provoke more intense experiences? And what exactly is the mapping function: is it linear, logarithmic, or a power function? And why is it one of these rather than another? Even these questions leave open the more fundamental question of how a neural code could ever give rise to experience at all.

Not very much scientific investigation has addressed this kind of question. Most scientists seem satisfied with some variant of Müller's (1838) classic concept of "specific nerve energy." Müller's idea, in its modern form,[3] amounts to the claim that what determines the particularly visual aspect of visual sensations is the fact that these sensations are transmitted by specific nerve pathways (namely, those originating in the retina and not in the cochlea) that project to particular cerebral regions (essentially, cortical area V1). It is certainly true that retinal influx comes together in relatively circumscribed areas of the brain, and that this may provide an architectural advantage in the neural implementation of the calculations necessary to generate visual-type sensations. But what is it about these pathways that generates the different sensations? Surely the choice of a particular subset of neurons or particular cortical regions cannot, *in itself*, explain why we attribute visual rather than auditory qualities to this influx. We could suppose that the neurons involved are of a different kind, with, say, different neurotransmitters, but then why and how do different neurotransmitters give rise to different experiences? We could say that the type of calculation done in the different cortical areas is different, but then we must ask, how could calculations ever give rise to experience? The hard work is left undone. Much still needs to be explained.

1.3. An alternative approach: The sensorimotor contingency theory

The present paper seeks to overcome the difficulties described above by adopting a different approach to the problem of visual experience. Instead of assuming that vision consists in the creation of an internal representation of the outside world whose activation somehow generates visual experience, we propose to treat vision as an *exploratory activity.* We then examine what this activity actually consists in. The central idea of our new approach is that *vision is a mode of exploration of the world that is mediated by knowledge of what we call sensorimotor contingencies.* We show that problems about the nature of visual consciousness, the qualitative character of visual experience, and the difference between vision and other sensory modalities, can now, from the new standpoint, all be approached in a natural way, without appealing to mysterious or arcane explanatory devices.

2. The structure of vision

As stated above, we propose that *vision is a mode of exploration of the world that is mediated by knowledge, on the part of the perceiver, of what we call sensorimotor contingencies.* We now explore this claim in detail.

2.1. Sensorimotor contingencies induced by the visual apparatus

Imagine a team of engineers operating a remote-controlled underwater vessel exploring the remains of the Titanic, and imagine a villainous aquatic monster that has interfered with the control cable by mixing up the connections to and from the underwater cameras, sonar equipment, robot arms, actuators, and sensors. What appears on the many screens, lights, and dials, no longer makes any sense, and the actuators no longer have their usual functions. What can the engineers do to save the situation? By observing the *structure of the changes* on the control panel that occur when they press various buttons and levers, the engineers should be able to deduce which buttons control which kind of motion of the vehicle, and which lights correspond to information deriving from the sensors mounted outside the vessel, which indicators correspond to sensors on the vessel's tentacles, and so on.

There is an analogy to be drawn between this example and the situation faced by the brain. From the point of view of the brain, there is nothing that in itself differentiates nervous influx coming from retinal, haptic, proprioceptive, olfactory, and other senses, and there is nothing to discriminate motor neurons that are connected to extraocular muscles, skeletal muscles, or any other structures. Even if the size, the shape, the firing patterns, or the places where the neurons are localized in the cortex differ, this does not in itself confer them with any particular visual, olfactory, motor or other perceptual quality.

On the other hand, what *does* differentiate vision from, say, audition or touch, is the *structure of the rules* governing the sensory changes produced by various motor actions, that is, what we call the *sensorimotor contingencies* governing visual exploration. Because the sensorimotor contingencies within different sensory domains (vision, audition, smell, etc.) are subject to different (in)variance properties, the structure of the rules that govern perception in these different modalities will be different in each modality.

A first law distinguishing visual percepts from perception in other modalities is the fact that when the eyes rotate, the sensory stimulation on the retina shifts and distorts in a very particular way, determined by the size of the eye movement, the spherical shape of the retina, and the nature of the ocular optics. In particular, as the eye moves, contours shift and the curvature of lines changes. For example, as shown in Figure 1, if you are looking at the midpoint of a horizontal line, the line will trace out a great arc on the inside of your eyeball. If you now switch your fixation point upwards, the curvature of the line will change; represented on a flattened-out retina, the line would now be curved. In general, straight lines on the retina distort dramatically as the eyes move, somewhat like an image in a distorting mirror.

Similarly, because of the difference in sampling density of the retinal photoreceptors in central and in peripheral vision, the distribution of information sensed by the retina changes drastically, but in a lawful way, as the eye moves. When the line is looked at directly, the cortical representation of the straight line is fat in the middle and tapers off to the ends. But when the eye moves off the line, the cortical representation peters out into a meager, banana-like shape, and the information about color is radically undersampled, as shown in the bottom right hand panel of Figure 1. Another law that characterizes the sensorimotor contingencies that are particular to visual percepts is the fact that the flow pattern on the retina is an expanding flow when the body moves forwards, and contracting when the body moves backwards. Visual percepts also share the fact that when the eyes close during blinks, the stimulation changes drastically, becoming uniform (i.e., the retinal image goes blank).

In contrast to all these typically *visual* sensorimotor contingencies, *auditory* sensorimotor contingencies have a different structure They are not, for example, affected by eye movements or blinks. They are affected in special ways by head movements: rotations of the head generally change the temporal asynchrony between left and right ears. Movement of the head in the direction of the sound source mainly affects the amplitude but not the frequency of the sensory input.

We therefore suggest that a crucial fact about vision is that visual exploration obeys certain laws of sensorimotor

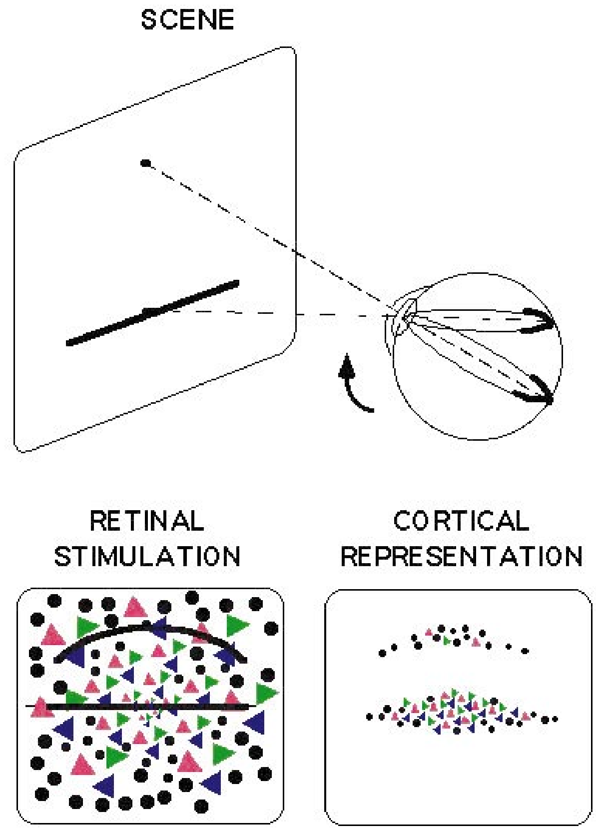

Figure 1. **Top:** The eye fixates the middle of a straight line and then moves to a point above the line. The retinal stimulation moves from a great arc on the equator of the eye to a different, smaller great arc. **Bottom left**: Flattened out retina showing great arc corresponding to equator (straight line) and off-equator great arc (curved line). Triangles symbolize color-sensitive cone photoreceptors, discs represent rod photoreceptors. Size of photoreceptors increases with eccentricity from the center of the retina. **Bottom right**: Cortical activation corresponding to stimulation by the two lines, showing how activation corresponding to a directly fixated straight line (large central oblong packet tapering off towards its ends) distorts into a thinner, banana shaped region, sampled mainly by rods, when the eye moves upwards. As explained in Section 2.2, if the eye moves *along* the straight line instead of upwards, there would be virtually *no change at all* in the cortical representation. This would be true even if the cortical representation were completely scrambled. This is the idea underlying the theory that shape in the world can be sensed by the laws obeyed by sensorimotor contingencies.

contingency. These laws are determined by the fact that the exploration is being done by the visual apparatus.

In summary: the sensorimotor contingencies discussed in this section are related to the visual apparatus and to the way three-dimensional objects present themselves to the visual apparatus. These sensorimotor contingencies are distinctive of the visual sense modality, and differ from the sensorimotor contingencies associated with other senses.

2.2. Sensorimotor contingencies determined by visual attributes

Real objects have properties such as size, shape, texture, and color, and they can be positioned in the three-dimensional world at different distances and angles with respect to an observer. Visual exploration provides ways of sampling these properties which differs from sampling via other senses. What characterizes the *visual* mode of sampling object properties are such facts as that the retinal image of an object only provides a view of the front of an object, and that when we move around it, parts appear and disappear from view; and that we can only apprehend an object from a definite distance, so that its retinal projection has a certain size that depends on distance. Other characteristics of visual exploration of objects derive from the fact that color and brightness of the light reflected from an object change in lawful ways as the object or the light source or the observer move around, or as the characteristics of the ambient light change.

On the other hand, tactile exploration of an object, even though it may be sampling the same objective properties, obeys different sensorimotor contingencies: you do not touch an object from a "point of view" – your hand can often encompass it more or less completely for example, and you don't apprehend it from different distances; its tactile aspect does not change with lighting conditions.

There is thus a subset of the sensorimotor contingencies that are engendered by the constraints of visual-type exploration, and which corresponds to *visual* attributes of sensed objects.

Note that unlike the sensorimotor contingencies that are visual-modality related, the sensorimotor contingencies that are visual-attribute related do, nonetheless, have strong links to the tactile sense: this is because attributes of three dimensional objects can also sometimes be apprehended via the tactile exploratory mode, where they present themselves as tactile shape, texture, size, distance. As shown eloquently by Piaget's work, the observer's conception of space in general will also have strong links to the laws of sensorimotor contingency discussed in the present section. Similar ideas were developed by Poincaré who wrote:

> To localize an object simply means to represent to oneself the movements that would be necessary to reach it. It is not a question of representing the movements themselves in space, but solely of representing to oneself the muscular sensations which accompany these movements and which do not presuppose the existence of space. (Poincaré 1905, p. 47)

A good illustration of sensorimotor contingencies associated with one particular kind of visual attribute, namely, visual shape, can be obtained from the records of patients whose vision has been restored after having been born blind with congenital cataract (cf. reviews by Gregory 1973; Jeannerod 1975; Morgan 1977). One such patient, cited by Helmholtz (1909/1925), is surprised that a coin, which is round, should so drastically change its shape when it is rotated (becoming elliptical in projection). The fact that objects also drastically change in extent as a function of distance is poignantly illustrated by the case of a 13–14 year old boy treated by Cheseldon (1728; cited by Morgan 1977, p. 20):

> Being shewn his father's picture in a locket at his mother's watch, and told what it was, he acknowledged a likeness, but was vastly surpriz'd; asking, how it could be, that a large face could be express'd in so little room, saying, it should have seem'd as impossible to him, as to put a bushel of any thing into a pint.

These examples make us realize how second nature it is for people with normal vision to witness the perspective changes that surfaces undergo when they are shifted or tilted, or when we move with respect to them. The idea we wish to suggest here is that the visual quality of shape *is precisely* the set of all potential distortions that the shape undergoes when it is moved relative to us, or when we move relative to it. Although this is an infinite set, the brain can abstract from this set a series of laws, and it is this set of laws which codes shape.[4]

Another example of how sensorimotor contingencies can be used as indicators of visual attributes is illustrated in an aspect of Figure 1 we have not yet mentioned. We saw in the introduction that movement of the eye away from a line creates a very strong distortion in its cortical and retinal representation. Under the classical view of what shape perception requires, it would be necessary to postulate that in order to see lines as straight despite eye movements, a transformation mechanism would have to exist that compensates for these distortions. This mechanism would take the cortical representation illustrated in the bottom right of the figure, and transform it so the two dissimilar packets of stimulated neurons shown in the figure now look identical.[5] There would additionally have to be another cortical locus where this new, corrected representation was projected. The view presented here does away with these unnecessary steps.

Consider the following fact: if the eye moves *along* the straight line instead of perpendicular to it, the set of photoreceptors on the retina which are stimulated does not change, since each photoreceptor that was on the image of the line before the eye moves is still on the image after the eye moves. This is due to an essential property of lines – they are self-similar under translation along their length (we assume, for simplicity, that the line is infinite in length). Since exactly the same photoreceptors are being stimulated before and after eye movement along the line's length, the cortical representation of the straight line is therefore identical after such a movement: there is this time no distortion at all. Another interesting fact is that the argument we have just made is totally independent of the code used by the brain to represent the straight line. Even if the optic nerve had been scrambled arbitrarily, or if the retina were corrugated instead of spherical, thereby causing the image of the line to be wiggly instead of straight, or if the eye's optics gave rise to horrendous distortions, movement of the eye along the line would still not change the pattern of cortical stimulation. We see that this particular law of sensorimotor invariance is therefore an intrinsic property of straight lines, and is independent of the code used to represent them. Platt (1960) has extended such considerations to other geometrical invariants, and Koenderink (1984a) has considered the more general, but related problem of how spatiotemporal contingencies in the neural input can be used to deduce intrinsic geometrical properties independently of the code by which they are represented.

In general, it will be the case that the *structure of the laws* abstracted from the sensorimotor contingencies associated with flat, concave, and convex surfaces, corners, and so on, will be a neural-code-independent indication of their different natures. In relation to this, some psychophysical work is being done; for example, to determine the respective importance, in determining shape, of cues derived from changes caused by movement of the object versus movement of the observer (e.g., Cornilleau-Peres & Droulez 1994; Dijkstra et al. 1995; Rogers & Graham 1979; Rogers & Rogers 1992). Nonetheless, though it is inherent in the approaches of a number of researchers (cf. sect. 3.3), the idea that the laws of sensorimotor contingency might actually *constitute* the way the brain codes visual attributes has not so far been greatly developed in the literature. However, this idea is essential in the present theory.

2.3. Sensation and perception

Psychologists interested in perception have traditionally distinguished between sensation and perception. While it is difficult to make this distinction precise, perhaps its central point is to differentiate between the way the senses are affected by stimuli (sensation) and the results of categorization of objects and events in the environment (perception). It is worthwhile to note that our distinction between two different classes of sensorimotor contingency roughly corresponds to this distinction between sensation and perception. Sensorimotor contingencies of the first sort – those that are determined by the character of the visual apparatus itself – are independent of any categorization or interpretation of objects and can thus be considered to be a fundamental, underlying aspect of visual *sensation.* Sensorimotor contingencies of the second sort – those pertaining to visual attributes – are the basis of visual *perception.*

In this way we can interpret the present theory as attempting to do justice to one of the working doctrines of traditional visual theory.

2.4. Perceivers must have mastery of patterns of sensorimotor contingency

Consider a missile guidance system allowing a missile to home in on an enemy airplane. As the missile zigzags around to evade enemy fire, the image of the target airplane shifts in the missile's sights. If the missile turns left, then the image of the target shifts to the right. If the missile slows down, the size of the image of the airplane decreases in a predictable way. The missile guidance system must adequately interpret and adapt to such changes in order to track the target airplane efficiently. In other words, the missile guidance system is "tuned to" the sensorimotor contingencies that govern airplane tracking. It "knows all about" or "has mastery over" the possible input/output relationships that occur during airplane tracking.

Now consider what happens when the missile guidance system is out of order. The visual information is being sampled by its camera, it is getting into the system, being registered, but it is not being properly made use of. The missile guidance system no longer has mastery over airplane tracking.

We suggest that vision requires the satisfaction of two basic conditions. First, the animal must be exploring the environment in a manner that is governed by the two main kinds of sensorimotor contingencies (those fixed by the visual apparatus, and those fixed by the character of objects). Second, the animal, or its brain, must be "tuned to" these laws of sensorimotor contingencies. That is, the animal must be *actively exercising* its mastery of these laws.

Note that the notion of being tuned, or having mastery, only makes sense within the context of the behavior and purpose of the system or individual in its habitual setting. Consider again the missile guidance system. If exactly the same system was being used for a different purpose, say, for example, as an attraction in a fun fair, it might well be necessary for the system to have a different behavior, with scary lunges and strong acceleration and deceleration which would be avoided in a real system. Thus, "mastery" of the sensorimotor contingencies might now require a different set of laws.[6] In fact even the out-of-order missile guidance system has a kind of ineffectual mastery of its sensorimotor contingencies.

2.5. Important upshot: A sensory modality is a mode of exploration mediated by distinctive sensorimotor contingencies

The present view is able to provide an account of the nature and difference among sensory modalities. In the introduction we stressed the deficiencies of Müller's (1838) view as well as of its modern adaptation,[7] according to which it is supposed that what determines the differences between the senses is some inherent characteristic of the neural pathways that are involved: this view requires postulating some special extra property which differentiates the neural substrate of these pathways, or some special additional mechanism, whose nature then stands in need of further (and for now at least unavailable) explanation. The present approach obviates this difficulty by saying that what differentiates the senses are the *laws* obeyed by the sensorimotor contingencies associated with these senses.[8] Hearing and audition are both forms of exploratory activity, but each is governed by different laws of sensorimotor contingency. Just as it is not necessary to postulate an intrinsic "essence" of horseriding to explain why it feels different from motorcycling, it is similarly unnecessary to postulate a Müller-type specific nerve energy to account for the difference between vision and other senses.[9]

The sensory modalities, according to the present proposal, are constituted by distinct patterns of sensorimotor contingency. *Visual perception can now be understood as the activity of exploring the environment in ways mediated by knowledge of the relevant sensorimotor contingencies.* And to be a visual perceiver is, thus, to be capable of exercising mastery of vision-related rules of sensorimotor contingency.

We shall see that this approach, in which vision is considered to be a law-governed mode of encounter with the environment, opens up new ways of thinking about phenomena such as synesthesia, the facial vision of the blind, and, in particular, tactile visual sensory substitution, where apparently *visual* experience can be obtained through arrays of vibrators on the skin.

2.6. Visual awareness: Integrating sensorimotor contingencies with reasoning and action-guidance

Thus far we have considered two important aspects of vision: the distinctively *visual* qualities that are determined by the character of the sensorimotor contingencies set up by the visual apparatus; and the aspect which corresponds to the encounter with visual *attributes,* that is, those features which allow objects to be distinguished visually from one another. These two aspects go some way towards characterizing the qualitative nature of vision.

We now turn to a third important aspect of vision, namely, *visual awareness.*

Suppose you are driving your car and at the same time talking to a friend. As you talk, the vista in front of you is impinging upon your eyes. The sky is blue, the car ahead of you is red, there is oncoming traffic, and so on. Your brain is tuned to the sensorimotor contingencies related to these aspects of the visual scene. In addition, some of these sensorimotor contingencies are also being used to control your driving behavior, since you are continuously adjusting your steering and adapting your speed to the moment-to-moment changes in the road and the traffic. But, since you are talking to your friend, you do not attend to most of these things. You do not notice that the car ahead is red, you do not think about the sky being blue; you just drive and talk to your friend.

You lack, as we shall say, *visual awareness* of many of the aspects of the visual scene. For those scene aspects, you are no different from an automatic pilot controlling the flight of an airplane. Your behavior is regulated by the appropriate sensorimotor contingencies, but you remain visually unaware of the associated aspects of the scene.

But if you should turn your attention to the color of the car ahead of you, and think about it, or discuss it with your friend, or use the knowledge of the car's color to influence decisions you are making, then, we would say, you are aware of it. For a creature (or a machine for that matter) to possess visual awareness, what is required is that, in addition to exercising the mastery of the relevant sensorimotor contingencies, it must make use of this exercise for the purposes of thought and planning.[10]

When you not only visually track an environmental feature by exercising your knowledge of the relevant sensorimotor contingencies, but in addition integrate this exercise of mastery of sensorimotor contingencies with capacities for thought and action-guidance, then you are visually aware of the relevant feature. Then, we say, you *see* it.

Consider an important point about this view of what visual awareness is, namely that our possession of it is a *matter of degree.* In particular, in our view, all seeing involves some degree of awareness, and some degree of unawareness. For example, if you were to probe an *unaware* driver waiting at the light, there would probably be some aspects of the red light that were at least indirectly being integrated into the driver's current action-guidance, rational reflection and speech. Perhaps, though not noticing the light's *redness,* the *fact* that the light was red may make him realize that he was going to be late. Or, though not noting that the light was red, the driver could be noting that it was difficult to see because the sky was too bright. On the other hand, even the driver who was *aware* of seeing the red light may not have been aware of all its aspects, for example, that the shape of the light was different from usual. A visual stimulus has a very large (perhaps infinite) number of attributes, and only a small number can at any moment be influencing one's action-guidance, rational reflection, and speech behavior.

A further important fact about this account of visual awareness is that it treats awareness as something *non*magical. There is no need to suppose that awareness and seeing are produced by the admixture of some mysterious additional element. To see is to explore one's environment in a way that is mediated by one's mastery of sensorimotor contingencies, *and* to be making use of this mastery in one's planning, reasoning, and speech behavior.

2.7. Visual consciousness and experience: Forms of awareness

It may be argued that there is still something missing in the present account of vision, namely, an explanation of visual *consciousness,* or of the *phenomenal experience* of vision. Although there is a great deal of disagreement among philosophers about these notions, there is broad consensus, first, that seeing involves experience in the sense that there is something it is like to see, and second, that it is somehow mysterious how we can possibly explain this subjective character of experience, or, as it is sometimes put, the "raw feel" or the "qualia" of vision, in neural or other physical terms. Is there any reason to believe the sensorimotor contingency approach can succeed here where others have failed?

We will return to some of these issues in section 6 of this paper. For now, let us note that the present sensorimotor contingency framework would seem to allow for the explanation and clarification, and certainly, for the scientific study of a good deal of what makes for the subjective character of experience. Thus, one important dimension of what it is like to see is fixed by the fact that there is a lawful relation of dependence between visual stimulation and what we do, and this lawful relation is determined by the character of the visual apparatus. A second crucial feature that contributes to what it is like to see is the fact that objects, when explored visually, present themselves to us as provoking sensorimotor contingencies of certain typically visual kinds, corresponding to visual attributes such as color, shape, texture, size, hidden and visible parts. Together, these first two aspects of seeing, namely, the visual-apparatus-related sensorimotor contingencies and the visual-object-related sensorimotor contingencies, are what make vision *visual,* rather than, say, tactile or auditory. Once these two aspects are in place, the third aspect of seeing, namely, visual awareness, would seem to account *for just about all the rest* of what goes into making up the character of seeing. For, visual awareness is precisely the availability of the kinds of features and processes making up the first two aspects for the purposes of control, thought, and action.

As said, the question of visual experience and consciousness is extremely controversial, and we will defer further discussion of our view until section 6.

3. Refinements of the view

Vision, we argue, requires *knowledge* of sensorimotor contingencies. To avoid misunderstanding, it is necessary to discuss this claim in greater detail.

3.1. Knowledge of sensorimotor contingencies is a practical, not a propositional form of knowledge

Mastery of the structure of the rules is not something about which we (in general) possess propositional knowledge. For example, we are not able to describe all the changes that a convex surface should suffer, or the distortions that should occur, upon moving our eyes to all sorts of positions on the surface, or when we move or rotate it. Nevertheless, our brains have extracted such laws, and any deviation from the laws will cause the percept of the surface's shape to be modified. Thus, for example, our brains register the fact that the laws associated with normal seeing are not being obeyed when, for example, we put on a new pair of glasses with a different prescription: for a while, distortions are seen when the head moves (because eye movements provoke displacements of unusual amplitudes); or when we look into a fish tank (now moving the head produces unusual kinds of distortions), or dream or hallucinate (now blinking, for instance, has no effect). Our impression in such cases is that, then, something unusual is happening.

3.2. Mastery must be currently exercised

Another important condition that we need to impose for sensorimotor contingencies to properly characterize vision, is that the mastery of laws of sensorimotor contingency be exercised *now.* The reason we need this condition is the following.

Over the course of life, a person will have encountered myriad visual attributes and visual stimuli, and each of these will have particular sets of sensorimotor contingencies associated with it. Each such set will have been recorded and will be latent, potentially available for recall: the brain thus has mastery of all these sensorimotor sets. But when a particular attribute is currently being seen, then the particular sensorimotor contingencies associated with it are no longer latent, but are actualized, or being currently made use of. In the language of the missile guidance system: the system may have stored programs that are applicable to the task of following different kinds of planes with different speed and turning characteristics. All these programs are latent, and the system has mastery of them all. But it is only when the system is following a particular type of plane, that it invokes and follows the particular recipe for that plane.

Again: among all previously memorized action recipes that allow you to make lawful changes in sensory stimulation, only some are applicable at the present moment. The sets that are applicable *now* are characteristic of the visual attributes of the object you are looking at, and their being *currently exercised* constitutes the fact of your visually perceiving that object.

3.3. Historical note: Relation to other similar ideas

Consider the following analogy with haptic perception, suggested by MacKay (1962; 1967; 1973). Suppose you are a blind person holding a bottle with your hand. You have the feeling of holding a bottle, you feel the bottle. But what sensations do you really have? Without slight rubbing of the skin, tactile information is considerably reduced, and even temperature sensation will, through adaptation of the receptors, disappear after you have held the bottle for a while. In fact therefore, you may well have *very little* sensory stimulation coming from the bottle at the present instant. Yet, you actually have the feeling of "having a bottle in your hand" at this moment. This is because your brain is "tuned" to certain *potentialities:* if you were to slide your hand very slightly, a change would come about in the incoming sensory signals which is typical of the change associated with the smooth, sliding surface of glass. Furthermore, if you were to move your hand upwards, the size of what you are encompassing with your hand would diminish (because you are moving onto the bottle's neck), and if you were to move downwards, your tactile receptors would respond to the roughness coming from the transition from glass to the paper label.

MacKay suggests that *seeing* a bottle is an analogous state of affairs.[11] You have the impression of seeing a bottle if there is knowledge in your nervous system concerning a certain web of contingencies. For example, you have knowledge of the fact that if you move your eyes up towards the neck of the bottle, the sensory stimulation will change in a way typical of what happens when a narrower region of the bottle comes into foveal vision; you have knowledge expressing the fact that if you move your eyes downwards, the sensory stimulation will change in a way typical of what happens when the white label is fixated by central vision. Similarly, motions of an object created by manual manipulation can be part of what visually differentiates objects from one another. Unlike a bottle, an object like a pitcher with a handle can be rotated and the handle made to appear and disappear behind the body of the pitcher. It is the possibility of doing this which is indicative of the fact that this is a pitcher and not a bottle. The visual nature of pitchers involves the knowledge that there are things that can be done to them which make a protrusion (the handle) appear and disappear.

Ryle (1949/1990) has made similar points. He says of a person contemplating a thimble:

> Knowing how thimbles look, he is ready to anticipate, though he need not actually anticipate, how it will look, if he approaches it, or moves away from it; and when, without having executed any such anticipations, he does approach it, or move away from it, it looks as he was prepared for it to look. When the actual glimpses of it that he gets are got according to the thimble recipe, they satisfy his acquired expectation-propensities; and this is his espying the thimble. (p. 218)

Other authors have, over the last decades, expressed similar views. Hochberg (1968, p. 323), for example, in the context of his notion of schematic maps, refers to: "the program of possible samplings of an extended scene, and of contingent expectancies of what will be seen as a result of those samplings," and Sperry (1952) has the notion of "implicit preparation to respond." These ideas are also related to Neisser's (1976) perceptual cycle, to Noton and Stark's (1971) "scanpath" theory, and were also put forward in O'Regan (1992) in relation to the notion of the "world as an outside memory." Although, as noted by Wagemans and de Weert (1992), Gibson's notion of "affordance" (see Gibson 1982b; Kelso & Kay 1987; Turvey et al. 1981) is sometimes considered "mystical," it is undoubtedly strongly related to our present approach (on this, see Noë 2002). The importance of action in perception has been stressed by Paillard (1971; 1991) and Berthoz (1997). Similar notions have also been found useful in "active vision" robotics (Ballard et al. 1997; Brooks 1987; 1991). Thomas (1999), in an excellent review, has advocated an "active perception" approach to perception and visual imagery, which corresponds very closely to our second, object-related type of sensorimotor contingency.

Another related viewpoint is to be found in the work of Maturana and Varela (1987/1992). Maturana and Varela also emphasize the importance of sensorimotor coupling for understanding the structure of the animal's cognitive and perceptual capacities, as well as for understanding the organization of the nervous system. Varela et al. (1991) present an "enactive conception" of experience according to which experience is not something that occurs inside the animal, but is something the animal enacts as it explores the environment in which it is situated (see also Noë et al. 2000; Pessoa et al. 1998; Thompson 1995; Thompson et al. 1992). A related approach has been put forward by Järvilehto (1998a; 1998b; 1999; 2000), who, in a series of articles with an approach very similar to ours,[12] stresses that perception is activity of the whole organism-environment system.

All these views of what it is to see – particularly MacKay's and Ryle's, – are based on the same notion of sensorimotor contingency that is so central to the view we are proposing in the present article. MacKay's work, especially, was the main source of inspiration of our theory. However, it should be emphasized that our view contains several novel elements not to be found in the works of these authors.

The first point we have stressed is that there is an important distinction to be made between the two classes of sensorimotor contingencies, those which are particular to the visual apparatus, and those which are particular to the way objects occupy three-dimensional space and present themselves to the eye. Most of the researchers cited in the previous paragraphs have been concerned mainly with the sensorimotor contingencies associated with visual object attributes. An exception may be the case of Gibson, who considered the more apparatus-related sensorimotor contingencies in different terms. In any case, it seems to us that it is mainly, though not exclusively, through these latter contingencies that we can give a principled account of the qualitative differences in the *experienced phenomenology* of the different sensory modalities, thereby providing a more principled alternative to Müller's notion of "specific nerve energy."

A second innovative point in our approach will become more evident in section 6. We shall see that by taking the stance that the experience of vision is actually *constituted* by a mode of exploring the environment, we escape having to postulate magical mechanisms to instill experience into the brain.[13]

4. The world as an outside memory

4.1. The world as an outside memory

Under the present theory, visual experience does not arise because an internal representation of the world is activated in some brain area. On the contrary, visual experience is a mode of activity involving practical knowledge about currently possible behaviors and associated sensory consequences. Visual experience rests on know-how, the possession of skills.

Indeed, there is no "*re*"-presentation of the world inside the brain: the only pictorial or 3D version required is the real outside version. What *is* required, however, are methods for probing the outside world – and visual perception constitutes one mode via which it can be probed. The experience of seeing occurs when the outside world is being probed according to the visual mode, that is, when the knowledge being accumulated is of the three kinds described above, that are typical of the visual modality.

Thus, as argued in O'Regan (1992), it could be said that the outside world acts as an *external memory* that can be probed at will by the sensory apparatus.

To further clarify this, it is useful to make the relation with normal memory. You know many things about where you live. But as you sit in your office, you may not be thinking about them. If you should start doing so, you can conjure up in your mind all manner of things. Each thing can be thought about in detail, but meanwhile, the other things, though latent, are not being thought about. As you think about your kitchen, your bedroom is not in your mind, though you can cause it to come to mind by merely thinking about it. Remembering is casting one's awareness onto parts of latent memories.

Similarly, seeing is casting one's awareness onto aspects of the outside world made available by the visual apparatus. As you look at a visual scene, you can interrogate yourself about different aspects of the scene. As soon as you do so, each thing you ask yourself about springs into awareness, and is perceived – not because it enters into a cortical representation, but because knowledge is now available about how sensations will change when you move your eyes, or move the object. However, before you actually wonder about some aspect of the scene, although the information is "out there," and although you know you can obtain it by making the appropriate eye movement or attention shift, it is not currently available. It is not currently available for being visually "chewed upon" or "manipulated," and cannot at this moment be used to control judgments and utterances: the third, "awareness" aspect of seeing is missing. Thus, even though the image of the object is impinging upon your retina, and even if its aspects are being analyzed by the feature-extracting modules of your visual system, under the current theory of seeing we must say that the object is not actually being seen.

As will be described in section 5, this way of thinking about vision brings with it a number of consequences about some classic problems related to the apparent stability of the visual world despite eye movements, and to the problem of "filling-in" or compensating for "imperfections" of the visual apparatus such as the blind spot. It also provided the impetus for the change-blindness experiments described in section 5.10.

4.2. The impression of seeing everything

A rather counter-intuitive aspect of the world-as-outside-memory idea, and the associated notion that there is no picture-like internal representation of the outside world, is that, in a certain sense, only what is currently being processed is being "seen." How then, – if at any moment only a small fragment of the world is actually being seen, – could we ever have that strong subjective impression that we continually have of seeing "everything"?

As pointed out by Noë et al. (2000) and Noë (2001), this paradox is actually only apparent, and rests on a misunderstanding of what seeing really is. It is true that normal perceivers take themselves to be aware of a detailed environment. But what this means is that they perceive the environment surrounding them as detailed. It does not mean that they think that inside their brains there is a detailed copy of the environment. It is only those perceivers – and there are many scientists among them – who make the mistake of thinking that "seeing" consists of making such a copy, who are led to think there is a problem.

Another way of understanding why our visual phenomenology is of seeing everything in front of us, derives from the fact that since the slightest flick of the eye or attention allows any part of a visual scene to be processed at will, we have the feeling of immediate availability about the whole scene. In other words, despite the fact that we are only currently processing a small number of details of the scene, under the present definition of seeing, we really are seeing the *whole* scene.

Suppose you should ask yourself, "Am I currently consciously seeing *everything* there is to see in the scene?" How could you check that you were seeing everything? You would check by casting your attention on each element of the scene, and verify that you have the impression of consciously seeing it. But obviously as soon as you do cast your attention on something, you see it. Conclusion, you will always have the impression of consciously seeing everything, since everything you check on, you see. There is an interesting and unfortunate consequence of this: If for some reason you should not be able to mentally attend to some aspect of the scene, you will not be able to consciously see it. Some empirical examples of this are given in sections 5.10–5.12.

One could make the amusing analogy, referred to by Thomas (1999), of the refrigerator light. It seems to be always on. You open the refrigerator: it's on. You close the refrigerator, and then open it again to check, the light's still on. It seems like it's on all the time! Similarly, the visual field seems to be continually present, because the slightest flick of the eye, or of attention, renders it visible. Brooks (1991) has said that the world should be considered as its own best model, and Minsky (1988) has suggested the notion of "immanence illusion" in a similar vein.

4.3. Vividness through transients

In addition to the "slightest flick of attention" argument there is another, very important, factor which explains the particular vividness of the feeling we have of a rich external visual presence. The visual system is particularly sensitive to *visual transients* (Breitmeyer & Ganz 1976; Stelmach et al. 1984; Tolhurst 1975). When a visual transient occurs, an automatic, "alerting" or "attention-grabbing" mechanism appears to direct processing to the location[14] where the transient occurred (Theeuwes 1991; Yantis 1998). This means that should anything happen in the environment, we will generally consciously see it, since processing will be directed to it. This gives us the impression of "having tabs" on everything that might change, and so, of consciously seeing everything. Were there not the attention-grabbing mechanism, our visual impression would be more similar to the impression we have when we stand with our backs to a precipice: we keenly feel it is there, we know that we can turn and see more of the precipice, but the feeling of presence is much less vivid than when we are actually looking into the precipice. The knowledge of having tabs on any change that might occur in the visual field – the fact that we know any change will attract our attention, – is another thing that makes the "outside memory" providing vision different from other forms of memory. For, any change in the visual field is immediately visible to you; whereas if, say, a Latin noun drops out of your memory overnight, no whistle will blow to let you know!

4.4. Dreaming and mental imagery

It is often claimed that dreaming, or other types of mental imagery, provide a counterexample to our denial that the brain must represent what is seen. Since dreams and mental images are apparently pictorial in nature, this seems to show that we *are,* after all, capable of creating an internal iconic image. Penfield's classic observations (e.g., Penfield & Jasper 1954) of visual memories being created by stimulation of visual cortex might also be thought to indicate that there are internal pictorial representations.

It is easy to be misled by these arguments, which for some reason are peculiarly compelling. But it is important to appreciate that they are misleading. Whether dreams, hallucination, or normal vision are at stake, these arguments are another instance of the error of thinking that when we see things as picture-like (be it when we look at reality or when we have a dream), this must be because there is some kind of internal picture. But this is as misguided as the supposition that to see red, there must be red neurons in the brain. The supposed fact that things appear pictorial to us in no way requires there to be pictures in the head. Therefore, the fact that we dream, hallucinate, and imagine does not provide evidence in favor of the view that the brain contains pictures of the detailed environment.[15]

A corollary of this confusion about dreams and mental imagery is the idea, expressed by a number of authors (e.g., Farah 1989; Kosslyn 1994; Zeki 1993) that feedback from higher brain areas into the retinotopic cortical map of area V1 would be a good way of creating mental imagery. This argument is somewhat misleading. It could be taken to be based on the implicit assumption that mental imagery occurs because of activation in V1: the topographic, metric layout of V1 would make it a good candidate for the cortical areas that possess what Zeki (1993) has called an "experiential" quality – that is, the capacity to generate experience. But again, the metric quality of V1 cannot in any way be the cause for the metric quality of our experience. It is as though, in order to generate letters on one's screen, the computer had to have little letters floating around in its electronics somewhere.

There may also be a second confusion at work in the argument from dreaming that we are considering. We have already noted that it does not follow from the fact that dreams are pictorial that, when we dream, there are pictures in the head. But do we really have reason to believe that dreams are pictorial? People certainly do say that they are. But does this give us reason to believe it is so? Just as we have observed that the idea that seeing is pictorial reflects a kind of naïve phenomenology of vision, it may very well be that the claim that dreaming is pictorial is similarly ill-founded phenomenologically. Certainly it is not the case that when we dream, it is as if we were looking at pictures. A hallmark of dream-like experiences is the unstable and seemingly random character of dreamt detail. For example, the writing on the card is different every time you look at it in the dream.[16] This suggests that without the world to serve as its own external model, the visual system lacks the resources to hold an experienced world steady.

4.5. Seeing without eye movements

Under the theory presented here, seeing involves testing the changes that occur through eye, body, and attention movements. Seeing without such movements is, under the theory, a subspecies of seeing: an exception. This would appear to be a rather dissident claim, given that psychologists studying visual perception have devoted a significant part of their energy precisely to the use of tachistoscopic stimulus presentation techniques, where stimuli are displayed for times shorter than the saccadic latency period of about 150 msec required for an eye movement to occur. Indeed, the studies show that observers are perfectly able to see under these conditions. For example, Potter (1976), in now classic experiments, showed that observers could pick out a target picture in a series of pictures presented at rates as fast as one picture every 125 msec. Thorpe et al. (1996) refined Potter's technique and showed by using event-related EEG potentials, that 150 msec after a stimulus is presented, that is, without any eye movement occurring, there is already information available in the cortex allowing the presence of an animal in a picture to be ascertained.

But because highly familiar stimuli (like words or animals) are used in these experiments, observers may be making use of a few distinctive features available in the images in order to accomplish the task. As argued by Neisser (1976), it probably cannot be said that observers are "seeing" the pictures in the normal sense of the word. As an illustration, consider an experiment we performed in which observers were asked to learn to distinguish three previously unknown symbols resembling Chinese characters (Nazir & O'Regan 1990). These were presented under the control of a computer linked to an eye movement measuring device. In the experiment, conditions were arranged so that observers could contemplate each Chinese symbol with their eyes fixated at the middle of the symbol, but as soon as the eyes moved, the symbol would disappear. Observers found this procedure extremely disrupting and irritating, and, contrary to what happens when the eye is free to move, hundreds of trials were necessary before they were able to distinguish the symbols. Furthermore, once the task was learnt, observers often failed when asked to recognize the learnt patterns at a new retinal location, only as little as half a degree away from the learnt position. Schlingensiepen et al. (1986) also found that without eye movements, observers had difficulty distinguishing patterns composed of arrays of random black and white squares; and Atkinson et al. (1976) showed by using an after-image technique that it is impossible to count more than four dots that are fixed with respect to the retina: a rather surprising fact. In a task of counting assemblies of debris-like pixel clumps, Kowler and Steinman (1977) found that observers had difficulties when eye movements were not permitted.[17] Because the stimuli used in these experiments were well above the acuity limit, the results are not explicable by acuity drop-off in peripheral vision. Even though a portion of the results may be due to lateral interaction effects (e.g., Toet & Levi 1992), it seems clear that observers are not at ease when doing a recognition task where eye movements are prohibited. It is like tactually trying to recognize an object lain on your hand without manipulating it.

A further suggestion of the need for visual exploration concerns the phenomenon of fading that occurs when the retinal image is immobilized artificially by use of an optical stabilization device. Under these circumstances a variety of perceptual phenomena occur, ranging from loss of contrast, to fragmentation, to the visual field becoming gray or "blacker than black" (Ditchburn 1973; Gerrits 1967). A portion of these phenomena can undoubtedly be accounted for in terms of the temporal response of the first stages of the visual system. Kelly (1982), for example, has suggested that detectors sensitive to oriented lines such as those discovered by Hubel and Wiesel actually are silent unless the oriented line stimulation is temporally modulated. Laming (1986; 1988) has stressed that neural transmission of external stimulation is always differentially coupled, so that, for example, the response of the retina to static stimulation is weak, and temporal modulation is necessary for optimal response (see also Arend 1973; Gerrits 1978; Kelly 1981; Krauskopf 1963).

From the point of view of the present theory, these phenomena are compatible with the idea that sensing of the visual world is a dynamic probing process. It could be that even the presence of a static external stimulus is not registered by a static sensory input, but by the dynamic pattern of the inputs that would potentially be produced by changes in the sensor position.

4.6. Why we don't see behind ourselves, but we do see partially occluded objects

Consider objects behind you, or in a box on your desk. Though you *know* that turning around or opening the box will cause certain changes in your sensory stimulation, some of which are indeed visual in nature, you do not have the feeling of *seeing* things behind you or in the box. The reason is that while the objects are behind you or in the box, the knowledge you have does not include certain essential visual aspects, namely, the knowledge that, say, blinking or moving your eyes will modify the sensations in a way typical of things that you see.

On the other hand, closer to normal seeing, consider an object which is partially occluded by another object. As you move your head, previously occluded parts appear, and previously unoccluded parts may disappear behind the occluder. This ability to make parts of the occluded object appear and disappear is similar to the ability to make objects appear and disappear by blinking, or to make their retinal projections change by moving the eye towards and away from them. This kind of ability is typical of what it is to see, so, even though the object is partially occluded, you nevertheless have the impression of seeing it, or at least "almost" seeing it. Furthermore, if you suddenly close your eyes and ask yourself exactly how much of the object was actually visible just before you closed your eyes, you will not generally know, and indeed, as suggested by results of Intraub and Richardson (1989), you will generally think you saw more than you did (see Fig. 2). This demonstrates that seeing is not directly related to having a retinal image, but to being able to manipulate the retinal image.

5. Empirical data

5.1. Introduction

In this section we will lay out a number of empirical results which are related to the theory of visual experience we have sketched. Before beginning however, it should be stressed that the empirical data to be presented is not intended as a *test* of the theory in the everyday sense in which theories are tested in science. We are providing a general framework for the study of vision, and it is not possible to subject a general framework to direct verification. Our new framework provides scientists with new problems and it makes some old problems appear as non-problems (like the problem of visual stability despite eye movements, and the problem of filling in the blind spot – see below). The framework highlights links between previously unrelated research streams, and creates new lines of research (like the work on change blindness, which was initiated by the idea of "the world as

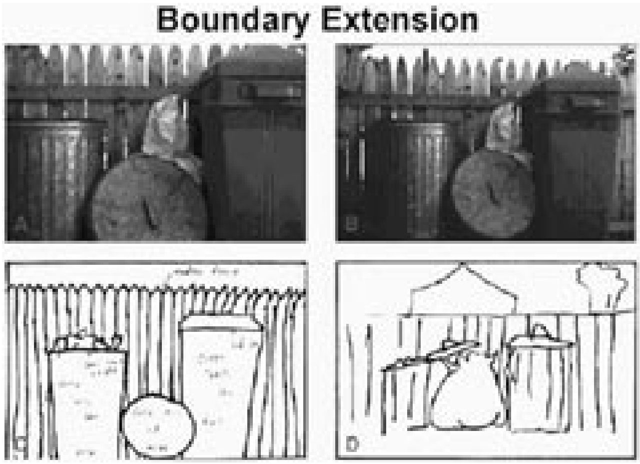

Figure 2. Subjects tend to remember having seen a greater expanse of a scene than was shown in a photograph. For example, when drawing the close-up view in Panel A from memory, one subject's drawing (Panel C) contained extended boundaries. Another subject, shown a more wide-angle view of the same scene (Panel B), also drew the scene with extended boundaries (Panel D). (Note: To evaluate the drawings in the figure, it is important to study the boundaries of each drawing and its associated stimulus.) (Figure and caption from Intraub, http://www.psych.udel.edu/~intraub)

an outside memory"). Of course, in each case, local, alternate, theories are possible within each of these domains, but the advantage of the present approach is that it brings together all the results so they can be seen from a single viewpoint.

In understanding the epistemological role of the present theory, an analogy can be made with the situation facing nineteenth-century physicists, who were trying to invent mechanisms by which gravitational or electrical forces could act instantaneously at a distance. To solve this problem, Faraday developed the idea of a *field of force* which was, according to Einstein, the single most important advance in physics since Newton (cf. Balibar 1992). But, in fact, the idea of a field of force is not a theory at all, it is just a new way of defining what is meant by force. It is a way of abandoning the problem being posed, rather than solving it. Einstein's abandoning the ether hypothesis is another example of how advances can be made by simply reformulating the questions one allows oneself to pose.

In the experiments to be described below, the first group relates to the notion that there is no picture-like internal representation of the outside world and that the world serves as an outside memory. These studies concern the problem of the apparent stability of the visual world despite eye movements, the filling-in of the blind spot and other (supposed) visual defects, and "change blindness": the fact that large changes in a visual scene sometimes go unnoticed. The second group of studies is more related to the idea that visual experience only occurs when there is the potential for action. These studies concern sensorimotor adaptation, sensory substitution, and synesthesia-related effects.

5.2. The extraretinal signal

At least since Helmholtz toward the end of the last century, a classic problem in vision has been to understand why the perturbations caused by eye movements (shift and smear on the retina) do not interfere with our perception of a stable visual world (cf. reviews of Bridgeman 1995; Grüsser 1986; MacKay 1973; Matin 1972; 1986; Shebilske 1977). A large portion of the experimental literature on the subject has assumed the existence of an internal representation, like a panoramic internal screen, into which successive snapshots of the visual world are inserted so as to create a fused global patchwork of the whole visual environment. The appropriate location to insert each successive snapshot is assumed to be determined by an "extraretinal signal," that is, a signal reflecting the direction the eyes are pointed at every moment. In total darkness some sort of extraretinal information is certainly available, as can easily be ascertained by noting that the after-image of a strong light source seems to move when the eyes move (Mack & Bachant 1969). Much debate has occurred concerning the question of whether the extraretinal signal is of efferent or afferent origin, and a convincing estimation of the role of the two components has been made by Bridgeman and Stark (1991). Irrespective of its origin however, the data concur to show that if the extraretinal signal exists, it is very inaccurate. Measurements from different sources (cf., e.g., compilations in Matin 1972; 1986) show that the signal must incorrectly be signaling that the eye starts to move as much as 200 msec before it actually does. The signal also incorrectly estimates the time and position where the eye lands, becoming accurate only about 1 second after the eye has reached its final position.[18] In any case, as admitted by Matin (1986), it is clear that the extraretinal information is too inaccurate, and also too sluggish, given the frequency of eye movements, to be used under normal viewing conditions to accurately place successive snapshots into a global fused internal image.

These results are not surprising when considered from the point of view of the theory of vision presented here. From this viewpoint, there is no need to postulate a mechanism that re-positions the retinal image after eye saccades so that the world appears stationary, because what is meant by "stationary," is precisely one particular kind of sensory change that occurs when the eye moves across an object. Having the feeling of seeing a stationary object consists in the knowledge that if you were to move your eye slightly leftwards, the object would shift one way on your retina, but if you were to move your eye rightwards, the object would shift the other way. The knowledge of all such potential movements and their results *constitute* the perception of stationarity. If on actually moving the eyes there were no corresponding retinal motion, the percept would not be of stationarity. From this point of view, there is no need to construct a stationary internal "image" of an object in order to see it as stationary. If there is such a thing as an internal signal in the brain that signals the eye's instantaneous position, then its purpose could not be to construct such an internal image (for there would be no one to look at it).

The question nevertheless arises of how the brain is able to accurately judge whether an object is stationary, or to control visuomanual coordination. If there is no way for retinal and extraretinal information to be combined to yield the true spatial coordinates of an object, how can the motion of an object ever be accurately ascertained, or how can an object be located with respect to the body and grasped? A possible answer may be that, whereas there is no extraretinal *signal,* there is nevertheless extraretinal *information* about the eye's location or velocity in the orbit. This information could be present in distributed form, and confounded with information about retinal stimulation. Such a distributed representation that mixes sensory and motor information (both of a static kind – position – and of a dynamic kind – velocity, acceleration) could provide the knowledge about sensorimotor contingencies required in the present theory. It could be used to perform accurate localization, but would not require the existence of a metric-preserving representation of the eye's position, or a picture-like internal image of objects on the retina or in space. Perhaps the multisensory neurons observed in parietal cortex, whose responses may be modulated by imminent eye movements, are compatible with this idea (Colby et al. 1996; Duhamel 1992; see also Zipser & Anderson 1988). Also of interest with respect to these ideas is a model of visual localization despite eye movements that has been constructed by Pouget and Sejnowski (1997). The model uses basis functions to code nonlinear mixtures of retinal and eye position. Linear combinations of these basis functions can provide pure retinal position, pure eye position, or head-centered coordinates of a target, despite the fact that no coherent internal map of the visual field has been constructed.

5.3. Trans-saccadic fusion

Over recent decades a new research topic has arisen with regard to the question of visual stability, in which researchers, instead of measuring the extraretinal signal itself, are questioning the notion that underlies it, namely, the notion of an internal screen in which successive snapshots are accumulated. The experimental methodology of this work consists in displaying stimuli which temporally straddle the eye saccade, and attempting to see if observers see a fused image – this would be predicted if an internal screen exists. Excellent reviews of this work (Irwin 1991; 1992) conclude that trans-saccadic fusion of this kind does not exist, or at least is restricted to a very small zone, namely, the zone corresponding to the target which the saccade is aiming for. Another kind of experiment consists in making large changes in high quality, full color pictures of natural scenes in such a way that the changes occur during an eye saccade (McConkie & Currie 1996). Even though the changes can occupy a considerable fraction of the field of view (e.g., cars appear or disappear in street scenes, swimming suits worn by foreground bathers change color, etc.), they are often not noticed – also contradicting the idea of a pictorial-type internal representation of the visual world. Again the conclusion appears to be that if there is an internal screen, it is not this internal screen which is providing us with the sensation of a stable, panoramic, visual world (Irwin & Andrews 1996; Irwin & Gordon 1998).

This conclusion is consistent with the theory presented here, where the problem of visual stability is a non-problem. Seeing does not require compensating for the effects produced by eye shifts in order to ensure accurate accumulation of partial views into a composite patchwork projected on some internal screen. There is no need to recreate another world inside the head in order for it to be seen. Instead, as suggested in section 4, the outside world acts as an "external memory" store, where information is available for probing by means of eye movements and shifts of attention O'Regan (1992).[19]

5.4. Saccadic suppression

Another issue which has preoccupied scientists concerns the question of why we are not aware of the smear caused by saccades. An enormous literature on the topic has been reviewed by Matin (1974; cf. also Li & Matin 1997): it appears that both at that time and still today (e.g., Burr et al. 1994; Li & Matin 1990; Ridder & Tomlinson 1997; Uchikawa & Sato 1995) many researchers believe that it is necessary to postulate some kind of suppression mechanism that inhibits transmission of sensory information to awareness during saccades, so that the rather drastic saccadic smear is not seen.

The empirical evidence showing diminished sensitivity to flashes during saccades cannot be denied, and the origin of this effect has been estimated by Li and Matin (1997) to be 20% due to the retinal smearing and masking caused by the image displacement (there may also be mechanical effects, as suggested by Richards 1969), and 80% due to central inhibitory mechanisms (some portion of this may be due to spatial uncertainty caused by the new eye position, cf. Greenhouse & Cohn 1991).

The important point, however, is that whatever inhibitory effects are occurring during saccades, these certainly do not constitute a suppression mechanism designed to prevent perception of the saccadic smear. If they did, then why would we not perceive a dimming of the world during saccades? Would we have to postulate a further un-dimming mechanism to compensate for the dimming? The notion of saccadic suppression probably constitutes another instance of the homunculus error, and is no less naive than postulating the need for a mechanism to right the upside-down retinal image so that the world appears right-side up. As explained in the theory presented above, there is no need to postulate mechanisms that compensate for the smear that is created by eye saccades, because this smear is part of what it is to see. If the retinal receptors did not signal a global smear during saccades, then the brain would have to assume that the observer was not seeing, and that he or she was perhaps hallucinating or dreaming.

5.5. Filling in the blind spot and perceptual completion

Another classic problem in vision which has recently been revived and generated heated debate (e.g., Ramachandran 1992; Ramachandran & Gregory 1991; Ramachandran 1995 vs. Durgin et al. 1995) is the problem of why we do not generally notice the 5–7 degree blind spot centered at about 17 degrees eccentricity in the temporal visual field of each eye, corresponding to the blind location on the retina where the optic nerve pierces through the eyeball.

Related problems involve understanding the apparent filling in of brightness or color that occurs in phenomena such as the Craik-O'Brian-Cornsweet effect and neon color spreading; the apparent generation of illusory contours as in the Kanisza triangle; and other phenomena of modal or amodal completion (cf. reviews of Kingdom & Moulden 1992; Pessoa et al. 1998).

Taking the case of the blind spot, from the point of view of the present theory, and in agreement with analyses of a number of theoreticians (Kingdom & Moulden 1992; Pessoa et al. 1998; Todorovič 1987), there is no need for there to be any filling in mechanism (O'Regan 1992). On the contrary, the blind spot can be *used* in order to see: if retinal sensation were not to change dramatically when an object falls into the blind spot, then the brain would have to conclude that the object was not being seen, but was being hallucinated. Suppose you explore your face with your hand: you can put your hand in such a way that your nose falls between two fingers. This does not give you the haptic impression of having no more nose. On the contrary, being able to put the nose between two fingers gives information about the size and nature of a nose. It is part of haptically perceiving the nose.

Monitoring the way the sensory stimulation from the retina changes when the eye moves to displace an object in the vicinity of the blind spot, is, for the brain, another way of gaining information about the object.

One can argue, however, that even though there may be no *need* for filling in processes, such filling in processes may nevertheless actually exist. In support of this, Pessoa et al. (1998), though critical of some neurophysiological and behavioral studies purporting to be evidence for filling in, concluded that several studies do point to the existence of precisely the kind of mechanisms which would be required for a filling in process. For example, Paradiso and Nakayama (1991), by using a masking paradigm, were able to measure the temporal dynamics of the phenomenal filling in of the inside of a bright disk. De Weerd et al. (1995) found cells in extrastriate cortex whose responses correlate well with the time it takes for the holes in textures presented in peripheral vision to perceptually fill in.

Just as was the case for the problem of the extraretinal signal or of saccadic suppression, the theory being advocated here does not deny the existence of neural mechanisms that underlie the perceptual phenomena that each of us observe. There can be no doubt that something is going on in the brain which is in relation to the fact that observers have no experience of a blind spot, and which makes Kanisza triangles have illusory contours. The question is: Is whatever is going on, actually serving to create an internal copy of the outside world, which has the metric properties of a picture, and which has to be completed in order for observers to have the phenomenology of a perfect scene? In the example of Paradiso and Nakayama's data, for example, there can be no denying that there must be retinal or cortical processes that involve some kind of dynamic spreading activation and inhibition, and that these processes underlie the percept that observers have in their paradigm – and possibly also when a disk is presented under normal conditions. But even though these processes *act like* filling in processes, this does not mean that they *are actually used by the brain* to fill in an internal metric picture of the world. They may just be providing information to the brain about the nature of the stimulation, but without this information being used to create a picture-like representation of the world.

In other words, our objection is not to the mechanisms themselves, whose existence we would not deny, but to the characterization of these mechanisms as involving "filling in." Consider this caricature: Spatio-temporal integration in the low-level visual system is a mechanism which explains much phenomenology (e.g., why fast flickering lights appear continuous, and very closely spaced dots look like lines). But surely no one would want to claim that the purpose of spatiotemporal integration is to "fill in" the temporal gaps in what would otherwise look like a stroboscopic world, or to make dotted lines look continuous. Spatiotemporal integration is a mechanism used in our visual systems to sample the environment, but its purpose is not to compensate for gaps in what would otherwise be a granular, pixel-like internal picture.

5.6. Other retinal non-homogeneities and the perception of color

A striking characteristic of the human visual system is its non-homogeneity. Spatial resolution is not constant across the retina, but falls off steadily: even the central foveal area is not a region of constant acuity, since at its edge (i.e., at an eccentricity of about 1 degree), position acuity has already dropped to half its value at the fovea's center (Levi et al. 1985; Yap et al. 1989). This drastic fall-off continues out into peripheral vision, only slowing down at around 15 degrees of eccentricity.

In addition to this non-homogeneity in spatial sampling, the retina also suffers from a non-homogeneity in the way it processes color: whereas in the macular region, the presence of three photoreceptor cone classes permits color discrimination, in the peripheral retina the cones become very sparse (Anderson et al. 1991; Coletta & Williams 1987; Marcos et al. 1996). The lack of the ability to accurately locate colors can easily be demonstrated by attempting to report the order of the colors of four or five previously unseen colored pencils when these are brought in from peripheral vision to a position just a few degrees to the side of one's fixation point.

A further, surprising non-homogeneity derives from the macular pigment, a yellowish jelly covering the macula, that absorbs up to 50% of the light in the short wavelength range (Bone et al. 1992), thereby profoundly altering color sensitivity in central vision.

Despite these non-homogeneities, the perception of spatial detail and color does not subjectively appear non-uniform to us: most people are completely unaware of how poor their acuity and their color perception are in peripheral vision. Analogously to the filling-in mechanism that is sometimes assumed to fill in the blind spot, one might be tempted to postulate some kind of compensation mechanism that would account for the perceived uniformity of the visual field. However, from the point of view of the present theory of visual experience, such compensation is unnecessary. This will be illustrated in relation to color perception below.

5.7. "Red" is knowing the structure of the changes that "red" causes

Under the present view of what seeing is, the visual experience of a red color patch depends on the *structure of the changes* in sensory input that occur when you move your eyes around relative to the patch, or when you move the patch around relative to yourself. For example, suppose you are looking directly at the red patch. Because of absorption by the macular pigment, the stimulation received by the color-sensitive retinal cones will have less energy in the short wavelengths when you look directly at the red patch, and more when you look away from the patch. Furthermore, since there is a difference in the distribution and the density of the different color-sensitive cones in central versus peripheral vision, with cone density dropping off considerably in the periphery, there will be a characteristic change in the relative stimulation coming from rods and cones that arises when your eyes move off the red patch. What determines the perceived color of the patch is the *set of such changes* that occur as you move your eyes over it.[20]

A relevant example arises from the perception of color in dichromats. When carefully tested in controlled conditions of illumination, dichromats exhibit deficiencies in their ability to distinguish colors, generally along the red-green dimension, which can be accounted for by assuming that they lack a particular type of cone, generally either the long or medium wavelength type. Curiously however, in real-life situations, dichromats are often quite good at making red-green distinctions. As suggested by Jameson and Hurvich (1978; cf. also Lillo et al. 1998) this is undoubtedly because they can make use of additional cues deriving from what they know about objects and what they can sense concerning ambient lighting. Thus, for example, when a surface is moved so that it reflects more yellowish sunlight and less bluish light from the sky, the particular way the spectrum of the reflected light changes, disambiguates the surface's color, and allows that color to be ascertained correctly even when the observer is a dichromat.

Though it is not surprising to find observers using all sorts of available cues to help them in their color discriminations, this kind of finding can be taken to support a much more far-reaching, fundamental hypothesis, put forward by Broackes (1992). This is that the color of a surface is not so much related to the spectrum of the reflected light, but rather, to the way the surface potentially *changes* the light *when the surface is moved with respect to the observer or the light sources.*

It must be stressed that more is being said here than was said by Jameson and Hurvich (1978), who merely noted that information is *available* that allows dichromats to make judgments similar to trichromats. Broackes' idea is that the colors of surfaces *are exactly* the laws governing the way the surface changes the reflected light.[21] At least as far as reflectivity of surfaces are concerned, the *same* laws apply to dichromats and trichromats, so that to a certain extent they have the *same* kinds of color perception: the difference is that dichromats have fewer clues to go by in many situations. Thus Broackes, who has color vision deficiencies[22] himself, claims that he has different experiences for red and green as do normals. His only problem is that sometimes, when lighting conditions are special, he can see certain dark red things as dark green, just as sometimes, in shadow, people with normal vision are convinced a garment is dark blue when in fact it is black, or vice versa. Of course, there will be a component of the sensorimotor contingencies, namely, those determined by the observer's own visual apparatus, which, to the extent that dichromats lack one of the three color channels, are different in the case of dichromats as compared to trichromats, so colors cannot be completely identical for them.

Broackes' theory of color is strongly related to the theory of visual perception that we have presented here. The difference between Broackes' views and ours is that Broackes is attempting to characterize the nature of *color* in terms of laws of sensorimotor contingency, whereas we have taken the bolder step of actually identifying color *experience* with the exercise of these laws, or, more precisely, with activity carried out in accord with the laws and based on knowledge of the laws.

5.8. Eye-position contingent perception

A surprising prediction from this idea, that the sensation of red comes from the *structure of changes* that is caused by red, is the following armchair experiment. Using a device to measure eye movements connected to a computer, it should be possible to arrange stimulation on a display screen so that whenever an observer looks *directly* at a patch of color it appears red, but whenever the observer's eye looks *away from* the patch, its color changes to green. The rather counterintuitive prediction from this is that after training in this situation, the observer should come to have the impression that green patches in peripheral vision and red patches in central vision are the same color.

Whereas exactly this kind of experiment has not yet been done, a variety of related manipulations were performed by McCollough (1965b) and by Kohler (1951).[23] For example, Kohler had observers wear spectacles in which one half of the visual field was tinted with blue, and the other half tinted with yellow. This is similar to the proposed armchair experiment in the sense that perceived color will be different depending on which way the observer moves the eyes. Results of the experiment seem to show that after adaptation, observers apparently came to see colors "normally." Similar phenomena were observed with half-prisms, in which the top and bottom portion of the visual field were shifted by several degrees with respect to each other. Observers ultimately adapted, so that manual localization of objects in the upper and lower visual fields was accurate.

Of particular interest in these studies would have been to know whether observers perceived the world as continuous despite the discontinuity imposed by the colored glasses or prisms. However, it is difficult to rigorously evaluate the reports, as they were only described informally by Kohler. Since then, though a large literature has developed over the last decades concerning many forms of perceptual adaptation, not very much work seems to have been done to investigate the effects of modifications like those imposed by the two-color glasses or the half-prisms, which produce strong discontinuities in the visual field.

Nevertheless, partial insight into such situations may be obtained by considering people who wear spectacles with bifocal lenses[24]: here a discontinuity exists in the visual field between the upper and lower part of the glasses. Depending on where an observer directs the eyes, the size and focus of objects will be different, because of the different power of the two parts of the lens. The question then is, does the world appear discontinuous to viewers of bifocals? The answer is that the *world* does not appear discontinuous, any more than the world appears "dirty" to someone who has not wiped his spectacle lenses clean. This is not to say that the observer cannot become aware of the discontinuity or the dirt on the lenses by attending to the appropriate aspect of the stimulation, just as it is possible to become aware of the blind spot in each eye by positioning a stimulus appropriately. But under normal circumstances the wearer of bifocals takes no notice of the discontinuity. Furthermore, even though image magnification as seen through the different parts of the lens are different, thereby modifying perception of distance, manual reaching for objects seen through the different parts of the lenses adapts and becomes accurate, as does the vestibulo-ocular reflex. Gauthier and Robinson (1975) and Gauthier (1976) have, for example, shown that wearers of normal spectacles with strong corrections, as well as scuba divers, come to possess a bistable state of adaptation, whereby their distance perception and reaching can instantaneously switch from one to the other state, as they take their spectacles on and off, or look through their underwater goggles (see also Welch et al. 1993 for a similar effect with prisms). In fact, an observer can be tricked into inappropriately switching adaptation state by surreptitiously removing the lenses from his or her eyeglasses, so that he or she incorrectly expects magnification to change when the eyeglasses are put on (Gauthier, personal communication).

5.9. Inversion of the visual world

Relevant to the theory of visual experience being proposed here, are the classic experiments performed by Stratton (1897), Kohler (1951), and some less often cited replications by Taylor (1962), Dolezal (1982), and Kottenhoff (1961), in which an observer wears an optical apparatus which inverts the retinal image so that the world appears upside-down and/or left-right inverted (cf. reviews by e.g., Harris 1965; 1980). Although at first totally incapacitated, observers adapt after a few days and are able to move around. Ultimately (after about two weeks of wearing the apparatus) they come to feel that their new visual world is "normal" again.[25]

What is interesting about these experiments is that during the course of adaptation, perception of the world is subject to a sort of fragmentation, and to a dependence on context and task. For example, Kohler (1951) reports that visual context allows something that is seen upside-down to be righted (e.g., a candle flips when it is lit because flames must go up, a cup flips when coffee is poured into it, because coffee must pour downwards). Ambiguities and inconsistencies abound: Dolezal reports sometimes being unable to prevent both his hands from moving when he tries to move only one. Kohler reports cases where two adjacent heads, one upright, the other inverted, were *both* perceived as upright. Kohler's observer Grill, after 18 days of wearing reversing spectacles, stands on the sidewalk and correctly sees vehicles driving on the "right," and hears the noise of the car motor coming from the correct direction. On the other hand, Grill nevertheless reports that the license plate numbers appear to be in mirror writing. Other observations are that a "3" is seen as in mirror writing, even though its open and closed sides are correctly localized as being on the left and right, respectively. The bicycle bell seems on the unusual side, even though the observer can turn the handle bars in the correct direction. Taylor (1962) has performed a study similar to Kohler's, except that instead of wearing the inverting spectacles continuously, his subject wore them only for a limited period each day. Under these conditions the subject rapidly obtains a bistable form of adaptation, adapted to both wearing and not wearing the spectacles. A point stressed by Taylor, in support of his behaviorist theory,[26] is that adaptation is specific to the particular body parts (arms, legs, torso) or activities (standing on both feet, on one foot, on the toes, riding a bicycle) that the subject has had training with, and that there is little "interpenetration" from one such sensorimotor system to another.

A theory of vision in which there is a picture-like internal representation of the outside world would not easily account for the fragmentation of visual perception described in these experiments: for example, it would be hard to explain the case of the license plate, where one aspect of a scene appears oriented accurately, and yet another aspect, *sharing the same retinal location,* appears inverted. On the other hand, the present theory, in which vision is knowledge of sensorimotor transformations, and the ability to act, readily provides an explanation: reading alphabetic characters involves a subspecies of behavior connected with reading, judging laterality involves another, independent, subspecies of behavior, namely, reaching. An observer adapting to an inverted world will in the course of adaptation only be able to progressively probe subsets of the sensorimotor contingencies that characterize his or her new visual world; and so inconsistencies and contradictions may easily arise between "islands" of visuo-motor behavior.[27]

Particularly interesting are cases of double vision when only one eye is open, that is, not explicable by diplopia. For example, Kohler's observer Grill saw two points of light when only one was presented slightly to the right of the median line (the second point was seen weaker, on the left, symmetrical to the original point). Similar observations of symmetrical "phantoms" were noticed by Stratton (1897), and can be compared to cases of monocular diplopia reported in strabismus (Ramachandran et al. 1994a; 1994b; Rozenblom & Kornyushina 1991). Taylor (1962) says of his subject wearing left-right inverting spectacles:

> Another of the training procedures he adopted was to walk round and round a chair or table, constantly touching it with his body, and frequently changing direction so as to bring both sides into action. It was during an exercise of this kind, on the eighth day of the experiment, that he had his first experience of perceiving an object in its true position. But it was a very strange experience, in that he perceived the chair as being both on the side where it was in contact with his body and on the opposite side. And by this he meant not just that he *knew* that the chair he saw on his left was actually on his right. He had that knowledge from the beginning of the experiment. The experience was more like the simultaneous perception of an object and its mirror image, although in this case the chair on the right was rather ghost-like. (pp. 201–202)

Presumably what happens in these experiments is that, because the spatial location or orientation of an object with respect to the body can be attributed either with respect to the pre- or the post-adapted frame of reference, during the course of adaptation it can sometimes be seen as being in both. Furthermore, orientation and localization of objects in the field of view can be defined with respect to multiple referents, and within different tasks, and each task may have adapted independently, thereby giving rise to incoherent visual impressions.

The impression we have of seeing a coherent world thus arises through the knitting together of a number of separate sensory and sensory-motor components, making use of visual, vestibular, tactile, and proprioceptive information; and in which different behaviors (e.g., reading, grasping, bicycle riding) constitute components that adapt independently, but each contribute to the experience of seeing. Conclusions of this kind have also been reached in a wealth of research on sensorimotor control, where it is shown that a gesture such as reaching for an object is composed of a number of sub-components (e.g., ballistic extension of the arm, fine control of the final approach and finger grasping, etc.), each of which may obey independent spatial and temporal constraints, and each of which may be controlled by different cerebral subsystems, which adapt separately to perturbations like changes in muscle proprioception, or in vestibular and visual information (for reviews of these results, see Jeannerod 1997; Rossetti et al. 1993).[28]

5.10. Change blindness experiments

The idea that the world constitutes an outside memory, and that we only see what we are currently attending to, was the impetus for a number of surprising experiments performed recently on "change blindness"[29] (O'Regan et al. 1999; 2000; Rensink et al. 1997; 2000). In these experiments, observers are shown displays of natural scenes, and asked to detect cyclically repeated changes, such as a large object shifting, changing color, or appearing and disappearing. Under normal circumstances a change of this type would create a transient signal in the visual system that would be detected by low-level visual mechanisms. This transient would exogenously attract attention to the location of the change, and the change would therefore be immediately seen.

However, in the change blindness experiments, conditions were arranged such that the transient that would normally occur was prevented from playing its attention-grabbing role. This could be done in several ways. One method consisted in superimposing a very brief global flicker over the whole visual field at the moment of the change. This global flicker served to swamp the local transient caused by the change, preventing attention from being attracted to it. A similar purpose could be achieved by making the change coincide with an eye saccade, an eye blink, or a film cut in a film sequence (for reviews, see Simons & Levin 1997).[30] In all these cases a brief global disturbance swamped the local transient and prevented it attracting attention to the location of the change. Another method used to prevent the local transient from operating in the normal fashion was to create a small number of additional, extraneous transients distributed over the picture, somewhat like mudsplashes on a car windscreen (cf. O'Regan et al. 1999). These local transients acted as decoys and made it likely that attention would be attracted to an incorrect location instead of going to the true change location.

The results of the experiments showed that in many cases observers have great difficulty seeing changes, even though the changes are very large, and occur in full view – they are perfectly visible to someone who knows what they are. Such results are surprising if one espouses the view that we should "see" everything that we are looking at: It is very troubling to be shown a picture where a change is occurring repetitively and in full view, without being able to see the change. The experience is quite contradictory with one's subjective impression of richness, of "seeing everything" in the visual field. However, the results are completely coherent with the view of seeing which is being defended here.

Another aspect of these experiments which relates to the present theory is a result observed in an experiment in which observers' eye movements were measured as they performed the task (O'Regan et al. 2000). It was found that in many cases, observers could be looking directly at the change at the moment the change occurred, and still not see it. Again, under the usual view that one should see what one is looking at, this is surprising. But under the view that what one sees is an aspect of the scene one is currently "visually manipulating," it is quite reasonable to observe that only a subset of scene elements that share a particular scene location should at a given moment be perceived.

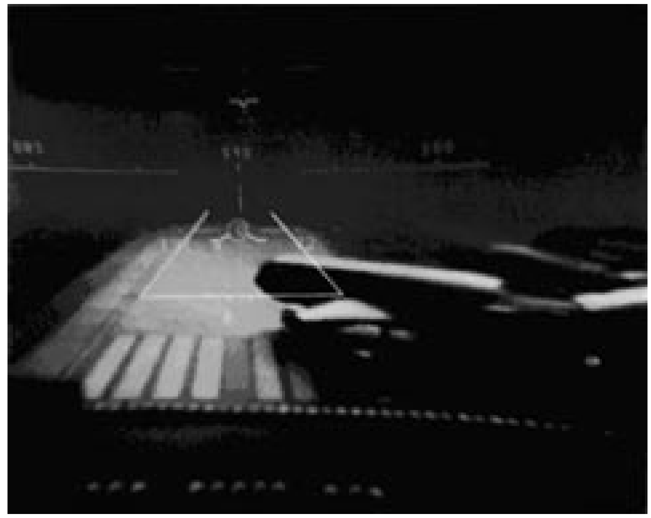

Figure 3. Simulator pilot's forward visual scene at an altitude of 72 feet and 131 knots with runway obstruction clearly visible. From Haines (1991). (Photo courtesy of NASA.)

A striking result of a similar nature had been observed by Haines (1991) and Fisher et al. (1980), who had professional pilots land an aircraft in a flight simulator under conditions of poor visibility, and using a head-up display (or "HUD") – that is, a display which superimposed flight guidance and control information on the windshield. On various occasions during the pilot's landing approach, they were presented with unexpected "critical" information in the form of a large jet airplane located directly ahead of them on the runway. Although the jet airplane was perfectly visible despite the head-up display (see Fig. 3), presumably because of the extreme improbability of such an occurrence, and because the pilots were concentrating on the head-up display or the landing maneuver, two of the eight experienced commercial pilots simply did not see the obstacle on the two occasions they were confronted with it, and simply landed their own aircraft through the obstacle. On later being confronted with a video of what had happened, they were incredulous.[31]

Other results showing that people can be looking directly at something without seeing it, had previously been obtained by Neisser and Becklen (1975), who used a situation which was a visual analogue of the "cocktail party" situation, where party-goers are able to attend to one of many superimposed voices. In their visual analogue, Neisser and Becklen visually superimposed two independent film sequences, and demonstrated that observers were able to single out and follow one of the sequences, while being oblivious of the other. Simons and Chabris (1999) have recently replicated and extended these effects.

Finally, Mack and Rock (1998) and Mack et al. (1992) have done a number of experiments using their paradigm of "inattentional blindness." In this, subjects will be engaged in an attention-intensive task such as determining which arm of a cross is longer. After a number of trials, an unexpected, perfectly visible, additional stimulus will appear near the cross. The authors observe that on many occasions this extraneous stimulus is simply not noticed.[32]

5.11. Inattentional amnesia

Related to the idea that the world serves as an outside memory, are the intriguing experiments of Wolfe (1997; 1999), and Wolfe et al. (1999) which they interpret in terms of what they call "inattentional amnesia."

Wolfe et al. (1999) use a standard visual search paradigm in which a subject must search for a target symbol among a number of distractor symbols. The authors estimate the efficiency of the search in milliseconds per item searched. However, instead of using a new display of distractors on each trial as is usually done, the authors use *exactly the same visual display* over a number of repetitions, but each time change the target that the subject is looking for. Since subjects are looking at the same display, which *remains continuously visible* on the screen for anything from 5 to 350 repetitions, depending on the experiment, one might have expected that an internal representation of the display would have time to build up, allowing search rate to improve over repetitions. However, this is not what is found: Over a number of experiments using different kinds of stimuli, Wolfe et al. (1999) find no evidence of improvement in search rate. It seems that no internal representation of the display is being built up over repetitions. In fact, search rate is as bad after many repeated searches as in the normal visual search conditions when the display changes at every trial: in other words, it is as though the subjects think they are searching through a brand new display at each trial, even though it is exactly the same display as before. Furthermore, an experiment done where the display is memorized and not visually presented at all, actually shows *faster* search speeds than when the display is present.

The results of these experiments are surprising under the view that what we see consists of an internal, more or less picture-like, representation of the visual world. However, they are exactly what would be expected under the present view, according to which "seeing" consists, not of having a "picture" in the mind, but of having seeking-out-routines that allow information to be obtained from the environment. Thus, observers generally do not bother to recreate within their minds a "re"-presentation of the outside world, because the outside world itself can serve as a memory for immediate probing. Indeed, the last result showing faster performance in the pure memory search shows that the very presence of a visual stimulus may actually obligatorily cause observers to make use of the world in the "outside memory" mode, even though it is less efficient than using "normal" memory.

This way of interpreting the results is also in broad agreement with Wolfe's point of view (Wolfe 1997; 1999) – Wolfe also refers to the notion of "outside memory." However, Wolfe lays additional emphasis on the role of attention in his experiments: Following the approach of Kahneman et al. (1992) adopted by many researchers in the attention literature, Wolfe believes that before attention is brought to bear on a particular region of the visual field, the elementary features (such as line segments, color patches, texture elements) analyzed automatically by low-level modules in the visual system constitute a sort of "primeval soup" or undifferentiated visual "stuff." Only once attention is applied to a particular spatial location, can the features be bound together so that an object (or recognizable visual entity) is perceived at that location. Wolfe's interesting proposition is that now, when visual attention subsequently moves on to another location, the previously bound-together visual entities disaggregate again and fall back into the "primeval soup": the previously perceived entity is no longer seen. This idea prompts Wolfe to use the term "inattentional amnesia," to emphasize the fact that after attention has moved on, nothing is left to see.

The status of the notion of attention in this explanation, and its relation to the theory presented here, is not entirely clear. One possibility would be to assume that what Wolfe means by "attention" is nothing other than visual awareness. In that case, the result of the experiment could be summarized by saying "once your awareness has moved off a part of the scene, you are no longer aware of it," which is tautological. Presumably, therefore, what Wolfe means by attention is something independent of awareness: there would be forms of attention without awareness and forms of awareness without attention. It is clear that further thought is needed to clarify these questions.

Independently of the framework within which one places oneself, it remains an interesting question to ask: What does the primeval soup "look like"? In other words, what does the visual field look like when the observer is not attending to anything in particular in it? Our preference would be to take the strict sense of attention in which attention = awareness, and to say that without attending to something (i.e., without being aware of anything), by *definition* the visual field cannot look like anything at all. Only when the observer attends to something will he or she be aware of seeing it. Note that what the observer attends to can be something as basic as overall brightness or color, or something like the variability in these ("colorfulness"?, "texturedness"?), or some attribute like "verticality" or "blobiness." If such features constitute the "primeval soup," then, like normal targets in the search task, the primeval soup would also only be "seen" if it was being attended to.

5.12. Informal examples

While the examples given in the preceding sections are striking experimental demonstrations of the fact that you do not always see where you look, several more informal demonstrations also speak to the issue.

Proofreading is notoriously difficult: when you look at words, you are processing words, not the letters that compose them. If there is an extra, incorrect letter in a word, it will have been processed by your low-level vision modules, but it will not have been "seen." Thus, for example, you will probably not have noticed that the "a"s in the last sentences were of a different shape than elsewhere.[33] Nonetheless on several occasions you were undoubtedly looking directly at them. It may take you a while to realize that the sign below (Fig. 4) does not say: The illusion of

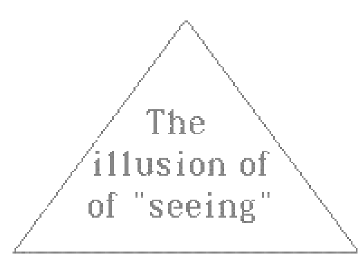

Figure 4. Ceci n'est pas: The illusion of "seeing."

Figure 5. Figure-ground competition.

"seeing."[34] You may be furious to find confirmation of years of the scientific study of reading showing that in this sentence there are in fact more "f"s than you think (count them!).[35]

The phenomena of figure-ground competition (see Fig. 5) and of ambiguous figures are also striking examples of how you do not see everything that you could see: when looking at such stimuli, you only see one of the possible configurations, even though more than one may be simultaneously available at the same location in your visual field.

It sometimes occurs that as you walk in the street you look directly at someone without seeing them. Only when the person gesticulates or manifests their irritation at not being recognized, do you become aware of who they are. While driving; it sometimes happens that you realize that you have been looking for a while at the brake lights of the car ahead of you without pressing on the brake.

5.13. Remote tactile sensing

An immediate consequence of the notion that experience derives not from sensation itself, but from the rules that govern action-related changes in sensory input, is the idea that visual experience should be obtainable via channels other than vision, provided that the brain extracts the same invariants from the structure of the sensori-motor contingencies.

A number of devices have been devised to allow people with deficits in one sensory modality to use another modality to gain information. In the domain of vision, two main classes of such sensory substitution devices have been constructed: echolocation devices and tactile visual substitution devices.

Echolocation devices provide auditory signals which depend on the direction, distance, size, and surface texture of nearby objects, but they provide no detailed shape information. Nevertheless, such devices have been extensively studied as prostheses for the blind, both in neonates (Bower 1977; Sampaio 1989; Sampaio & Dufier 1988) and in adults (Ifukube et al. 1991). It is clear that while such devices obviously cannot provide visual experience, they nevertheless provide users with the clear impression of things being "out in front of them."

Particularly interesting is the work being done by Lenay (1997), using an extreme simplification of the echolocation device, in which a blind or blindfolded person has a single photoelectric sensor attached to his or her forefinger, and can scan a simple environment (e.g., consisting of several isolated light sources) by pointing. Every time the photosensor points directly at a light source, the subject hears a beep or feels a vibration. Depending on whether the finger is moved laterally, or in an arc, the subject establishes different types of sensorimotor contingencies: lateral movement allows information about direction to be obtained, movement in an arc centered on the object gives information about depth. Note several interesting facts. First, users of such a device rapidly say that they do not notice vibrations on their skin or hear sounds, rather they "sense" the presence of objects outside of them. Note also that at a given moment during exploration of the environment, subjects may be receiving no beep or vibration whatsoever, and yet "feel" the presence of an object before them. In other words, the experience of perception derives from the *potential* to obtain changes in sensation, not from the sensations themselves. Note also that the exact nature or body location of the stimulation (beep or vibration) has no bearing on perception of the stimulus – the vibration can be applied on the finger or anywhere else on the body. This again shows that what is important is the sensorimotor invariance structure of the changes in sensation, not the sensation itself.

Lenay's very simple setup provides a concrete example of what is meant by the laws of sensorimotor contingency. Suppose that the photosensor were mounted on the forearm of an articulated arm, with the arm making an angle α with the torso, and the forearm making an angle β with the arm, as shown in the Figure 6. Then we can define the sensorimotor manifold as the two-dimensional space α: $[0, \pi/2]$ and β: $]3\pi/2\text{-a}, 2\pi[$. Consider the situation where we are obtaining information about depth by making movement in an arc. If a luminous source at distance L is being "fixated," the angles α and β will lie on orbits in the sensorimotor sensorimotor manifold defined by the relation shown in the lower part of the figure. In reality of course the angles α and β will be nonlinear functions of high-dimensional neural population vectors corresponding to arm and forearm muscle parameters. But the laws of contingency will be the same.

On further reflection it is apparent that the simple device studied by Lenay is an electronic variant of the blind person's cane. Blind persons using a cane do not sense the cane, but the outside environment that they are exploring by means of the cane. It has been said that the tactile sensations provided by the cane are somehow "relocated" or "projected" onto the environment. The cane itself is forgotten or ignored. But this way of describing experience with the cane, though in a way correct, is misleading, as it suggests that sensations themselves originally possessed a location which had to be relocated. The present theory shows that in themselves, sensations are situated nowhere. The location of a sensation (and, for that matter, any perceived aspect including its moment of occurrence) is an abstraction constructed in order to account for the invariance structure of the available sensorimotor contingencies.

Note that similar experiences to those of the blind person with the cane are experienced every day even by sighted persons: Car drivers "feel" the wheels on the road, and extend the sense of their bodies to include the whole car, allowing them to negotiate into parking spaces with only centimeters to spare. A particularly poignant example of having one's perceived body extend outside of the boundary formed by the skin was given to the first author by a friend who is a talented viola player. Spending most of the

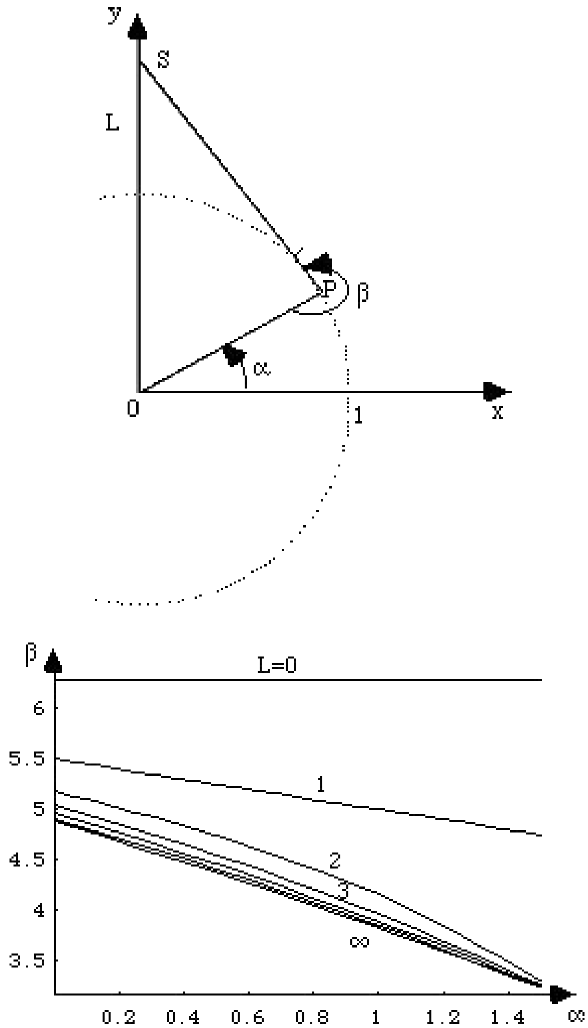

Figure 6. Figure from Lenay et al. (1997) showing the sensorimotor contingency for a simple photocell mounted on an arm and a forearm, in the case where the photocell is continuously fixating a luminous source at distance L. **A:** The arm (with the forearm) has a length of 1. The distance from the target, L (0S), can then be obtained by a trigonometrical relation, according to the following formula: (1) $L = \sin \alpha - \cos \alpha \tan (\alpha + \beta)$, where α: $[0, \pi/2]$ and β: $[3\pi/2 - a, 2\pi]$. **B:** Curve representing angle β in relation to angle α (both expressed in radians) for the following values of $L = 0,1, \ldots 7$. α varies from 0 to $\pi/2$. According to (1) one can determine β for any given L and α: $\beta = 2\pi - \alpha + \text{Atan}((\sin \alpha - L)/\cos \alpha)$.

day with the viola under his chin, on one occasion he went into the kitchen to drink some hot tea, and some drops fell on the viola. He said he was surprised not to have felt the hot drops on the instrument: it felt anesthetized. Another everyday example of remote tactile sensing occurs when you write on a piece of paper with a pen: you feel the paper at the end of the pen: it is rough, it is smooth, it is soft. You locate the contact at the end of the pen, not on your fingers where the force is actually felt (this example is given by James 1890/1950).

One might consider these examples as surprising at first sight. But then we ask: should it not also be considered surprising that fingertip sensations are felt *on the fingertips,* since after all, it is presumably in the brain where the sensations are registered? Why would one not tend to think that one should be able to walk through a door no wider than one's brain, since body sensations presumable arrive in the brain? Indeed, given that visual sensation impregnates the retina, why does one not feel the outside world as situated on one's retina, instead of outside one? These obviously ridiculous extensions of the "relocation" idea discussed above make one realize that, actually, the perceived location of a sensation cannot be logically determined by where the nerves come from or where they go to. Perceived location is, like other aspects of sensation, an abstraction that the brain has deduced from the structure of the sensorimotor contingencies that govern the sensation.[36]

Some very interesting experiments of Tastevin (1937) are related to these points. Tastevin had shown that the sensed identity or position of a limb can be transferred to another limb or to a plaster model of the limb. Thus, for example, when an experimenter feigns to touch a subject's forefinger with one prong of a compass, but actually touches the middle finger with the other prong, the subject feels the touch on the forefinger. Sensation has thus been relocated from the middle finger to the forefinger. Whole body parts can be relocalized by this means. A recent experiment along very similar lines was described by Botvinick and Cohen (1998; and also extended by Ramachandran & Blakeslee 1998). These authors used a life-size rubber model of a left arm placed before a subject whose real left arm was hidden by a screen. Using two small brushes, the experimenters synchronously stroked corresponding positions of the rubber and real arm. After ten minutes, subjects came to feel that the rubber arm was their own.

All these phenomena show how labile the perceived location of a stimulation can be, and how it depends on correlation with information from other modalities (in this case vision). Even neural representations of body parts are known to be labile, as has been shown by Iriki et al. (1996) whose macaque monkeys' bimodal visual somatosensory receptive fields moved from their hands to the ends of a rake they used as a tool. However, a facile interpretation of such phenomena in terms of "neural plasticity" of cortical maps would be misleading, since such an interpretation would implicitly assume that perceived location of a stimulus is directly related to activity in cortical maps – an idea we reject.

5.14. Tactile visual sensory substitution

Tactile visual substitution systems (TVSS) use an array of vibratory or electrical cutaneous stimulators to represent the luminance distribution captured by a TV camera on some skin area, such as the back, the abdomen, the forehead or the fingertip. For technical reasons and because of the restrictions on tactile acuity, TVSS devices have up to now suffered from very poor spatial resolution, generally having stimulator arrays of not more than 20×20 stimulators at the very best. They have also been bulky, expensive, and too sensitive to light level variations, for them to be of practical use by the blind (Bach-y-Rita 1983; Easton 1992). Notwithstanding these problems, however, as concerns the question of visual experience, a number of highly interesting points have been made about the experiences of individuals who have used these devices (Apkarian 1983; Guarniero 1977; 1974).

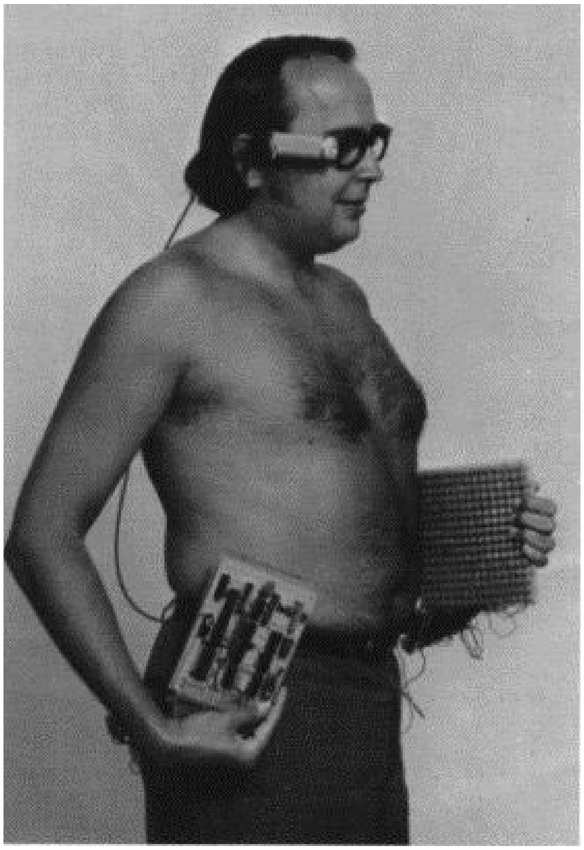

Figure 7. A blind subject with a "Tactile Visual Substitution system" (TVSS). A TV camera (mounted on spectacle frames) sends signals through electronic circuitry (displayed in right hand) to an array of small vibrators (left hand) which is strapped against the subject's skin. The pattern of tactile sitmulation corresponds roughly to a greatly enlarged visual image. (Photograph courtesy of P. Bach-y-Rita). From Morgan (1977).

A first point concerns the importance of the observer's being able to manipulate the TV camera himself or herself (Bach-y-Rita 1972; 1984; Sampaio 1995).

In the earliest trials with the TVSS device, blind subjects generally unsuccessfully attempted to identify objects that were placed in front of the camera, which was fixed. It was only when the observer was allowed to actively manipulate the camera that identification became possible and observers came to "see" objects as being externally localized (White et al. 1970). This important point constitutes an empirical verification of the mainstay of the present theory of visual experience, namely, that seeing constitutes the ability to actively modify sensory impressions in certain law-obeying ways.

Once observers have had practice with the TVSS, several further aspects of the experience provided by the system suggest that it is similar to the experience of vision. First, though initially observers locate stimulation on the body part which is stimulated, with practice, the observers locate objects in space, and not on the skin – although they are still able to feel the local tactile sensation (e.g., if it is painful or if it itches). Indeed, after using one skin location (e.g., the back), an observer has no problem transferring to a different skin location (e.g., the forehead).

An interesting example shows that the localization of objects outside the body is not just a cognitive strategy but truly resembles visual experience. In an anecdote reported by Bach-y-Rita, the zoom control of the camera being used by a well-trained subject was moved, causing a sudden magnification or "looming" of the tactile image. Bach-y-Rita states (1972, p. 98): "the startled subject raised his arms and threw his head backward to avoid the 'approaching' object. It is noteworthy that, although the stimulus array was, at the time, on the subject's back, he moved *backward* and raised his arms in front to avoid the object, which was subjectively located in the three-dimensional space before him."[37] Another interesting observation caused puzzlement in the early investigations with the TVSS. For practical reasons the battery of 400 vibrators mounted on the observer's back consisted of two ramps of 200 vibrators, one on each side of the observer's backbone. A large gap was therefore present in the tactile representation of the visual field. "Curiously" however, no gap was apparent in observers' perceived visual field. This tactile analog of what might incorrectly be called "filling-in" of the retinal blind spot is, of course, unsurprising in the light of the present theory, where no filling-in mechanism need be postulated (cf. sect. 5.5).

Do blind people actually see with the TVSS? The question has been raised by Bach-y-Rita who prefers to put the word "see" in quotes. One justification for this, he claims, is the fact that people who have learnt to see with the device are disappointed when shown pictures of their loved ones, or erotic pictures: they have no emotional reaction. Bach-y-Rita interprets this as a failure of the device to provide true visual experience. An alternative, however, is to admit that the device does provide true visual experience, but that emotional and sexual reactions are strongly linked to the sensations that are experienced during the period when emotional attachment occurs and sexual interest develops. If, during the course of development, these experiences are initially non-visual, then they will remain non-visual.[38]

Morgan (1977) also discusses this and concludes, that either people really do see with the TVSS, or there can be no scientific psychology. Clearly from the point of view of the present theory, seeing is not a matter of "all or nothing." There are many aspects to seeing, and the TVSS provides some but not all of them. The invariants related to position and size changes of the tactile image are similar to those in normal vision. Color and stereo vision however are absent, and resolution is extremely poor. But, just as color blind, stereo blind, one-eyed or low-sighted people can be said to "see," people using the TVSS should also be said to see. The fact that stimulation is provided through the skin should be irrelevant, providing the stimulation obeys the required sensorimotor laws. Of course, seeing with the skin probably involves laws that are not exactly the same as seeing with the eyes, just as seeing colors in the dark is not quite the same as in the light. The experience associated with the TVSS will thus also be somewhat different from normal visual experience.

5.15. The "facial vision" of the blind

A further interesting example of sensory substitution comes from what is called the "facial vision," or "obstacle sense," or "pressure sense" of blind people. In locating objects, particularly when these are large and in the 30–80 cm range, blind people often have the impression of a slight touch on their forehead, cheeks, and sometimes chest, as though they were being touched by a fine veil or cobweb (James 1890/1950; Kohler 1967).

For instance, consider the following quote given by James (1890/1950) from the blind author of a treatise on blindness of the time:

> Whether within a house or in the open air, whether walking or standing still, I can tell, although quite blind, when I am opposite an object, and can perceive whether it be tall or short, slender or bulky. I can also detect whether it be a solitary object or a continuous fence; whether it be a close fence or composed of open rails, and often whether it be a wooden fence, a brick or stone wall, or a quick-set hedge. . . . The currents of air can have nothing to do with this power, as the state of the wind does not directly affect it; the sense of hearing has nothing to do with it, as when snow lies thickly on the ground objects are more distinct, although the footfall cannot be heard. I seem to perceive objects through the skin of my face, and to have the impressions immediately transmitted to the brain. (Vol. 2, p. 204).

At least since Diderot's "Letter on the blind," facial vision had often been considered to truly be a kind of tactile, or even possibly an extrasensory, form of perception (cf. historical review by Hayes 1935, cited by Rice 1966). James (1890/1950, Vol 2, pp. 140, 204) compares this sense to what he believes is a tactile, pressure-related "tympanic sense," that is, the ability we all have of sensing with closed eyes whether an object brought before our face is large or small, or more or less solid. Despite such claims, however, by stopping up the ears of blind people with putty, James demonstrated to his satisfaction that audition was involved in the facial sense. This was then definitively established by Dallenbach et al. (1944), and facial vision is now known to be essentially caused by intensity, direction, and frequency shifts of reflected sounds (see review by Arias 1996). Kohler (1967) actually went so far as to anesthetize the faces of blind people, who nevertheless continued to have these sensations.

As noted by Worchel et al. (1950; cited by Strelow & Brabyn 1982), the question arises why this form of object perception is experienced as having a tactile feeling rather than an auditory quality. A possibility along behaviorist lines has been suggested by Taylor (1962), who supposes that collisions with obstacles will often involve the face – the hands may often rise and protect the face. This may create, by association, feelings on the face in the case of impeding collisions. Further correlations (apparently not mentioned by Taylor) might be the fact that objects that are close to the face tend to provoke slight disturbances of the air as well as changes in heat radiation that could be detected by receptors on the face. Although Taylor's associationist hypothesis may have some truth in it, from the point of view of the present theory another possibility arises: the prediction would be that the sensorimotor contingencies created by the particular, very subtle information received through the auditory modality would, in this particular case, have an invariance structure that resembles the contingencies caused by tactile stimuli, like those created by a veil upon the face.

Indeed, it appears conceivable that the object sense, requiring more subtle auditory distinctions, would be much more critically dependent on distance than normal hearing. In particular, moving a few centimeters forward or backwards, might create a radical change analogous to moving a few centimeters forward or backwards and bringing the head into and out of contact with a veil. Similarly, it may be that when the head is facing the object that is being sensed, slight sideways shifts of the head might create systematic changes similar to the systematic rubbing that occurs when one is touching a piece of cloth with the head. Note, however, that it would be exaggerated to take too literally the comparison that blind people make with veils and cobwebs: Kohler has verified that when touched with actual veils the same blind people say that the sensations are actually quite different. Perhaps the inability to specify precisely the nature of the experience produced prompted the author cited by James to say that the impressions were "immediately transmitted to the brain."

The facial sense of the blind may be related to the phenomenon of synesthesia (Cytowic & Wood 1982; Baron-Cohen & Harrison 1996),[39] where a stimulus in one sensory modality evokes sensations in another, the most frequently occurring case being colored hearing (Marks 1978). Ventriloquism is another type of example where information from one sensory modality modifies that in another: in "visual capture" or the "ventriloquism effect," the perceived location of a sound source is influenced by its seen position, and, to a lesser extent, vice versa (Hatwell 1986; Radeau & Bertelson 1974; Warren et al. 1981). A related phenomenon is the McGurk effect (McGurk & MacDonald 1976) in which the identity of a heard phoneme is altered by simultaneously observing a visual display of a different phoneme being pronounced. Radeau (1997; cf. also Marks 1978) has reviewed a number of inter-sensory interactions such as these, both in humans and animals, and concludes that such effects are compatible with the notion that the different qualities of the senses are not present ab initio, as Piaget might have claimed, but rather (following Gibson 1966; Bower 1979) are the result of a progressive differentiation process that occurs in the developing organism through the influence of environmental experience.

The view taken within the context of the present theory regarding all such intermodal interactions would be related to the above. More precisely, however, it would say that the experience associated with a modality exists only within the context of the acting organism, and within the context of the other senses available to the organism. Although vision, audition, touch, and so on, will have their own specificities due to the particularities of the sensors and sensorimotor contingencies involved – with these specificities defining the particular experience associated with each sense, – interactions between the senses are to be expected when there are systematic correlations and common sensorimotor contingencies.[40] Perceptual adaptation effects like the McCollough effect (Harris 1980; Humphrey et al. 1994; McCollough 1965a; 1965b) and the related disappearance of color fringes on adaptation to displacing prisms (Held 1980; Kohler 1951) may be manifestations of similar nature, except that they are intramodal rather than intermodal.

6. Visual consciousness

6.1. Introduction

The sensorimotor contingency theory of vision we have developed here provides a new vantage point from which to approach the vexing theoretical question of the nature of visual consciousness. Vision, we have argued, is a mode of skillful encounter with the environment, requiring knowledge of sensorimotor contingencies and also the ability to make use of that knowledge for the purpose of guiding action, thought, and (in humans) language use. What, in this view, is visual consciousness?

6.2. Two kinds of visual consciousness

We propose to distinguish between two kinds of visual consciousness: (1) *transitive visual consciousness* or *consciousness of;* and (2) visual consciousness in general.

(1) To be *transitively conscious* is to be conscious *of* a feature of a scene (Malcolm 1984). To be conscious of a feature of a scene in this sense is simply to be visually aware of it, as laid out in section 2.6.

Thus, to say that you are transitively conscious of (for example) the shape of a parked car in front of you is to say that you are, first, currently exercising mastery of the laws of sensorimotor contingency that pertain to information about the shape of the car; and, second, that you are attending to this exercise, in the sense that you are integrating it into your current planning, reasoning, and speech behavior.

Notice that when you are visually conscious of the shape of the car, you may fail to attend to its color, or to the fact that the object in front of you *is* a car. As you shift your attention from one aspect of the car to another, features of the car enter consciousness. What happens when you thus shift your attention is that you draw into play different bits of implicit knowledge of the relevant sensorimotor contingencies.

To this, it might be objected that when you look at the car, you have the impression that all its details are available in consciousness all at once. In looking at the car, you are conscious of, or aware of, its shape, color, nature, and so on, all in a glance. But this objection is easily countered. First, the empirical data on change blindness (see sect. 5.10 above) and inattentional blindness (see Noë & O'Regan 2000) demonstrate that you do *not* have all the details of what is seen in consciousness at once. The actual case is that all the detail is there in the scene in front of you and is thus accessible by the slightest effort. Second, and of great importance, is that your *feeling* of the presence of all the detail consists precisely in your knowledge that you can access all this information by movements and inquiries.

(2) *Visual consciousness in general,* on the other hand, is a higher-order capacity. To be visually conscious in general is to be poised to become aware of a present feature (that is, to become transitively conscious of it). In this sense of visual consciousness, we can contrast *being visually conscious* with being asleep or with being blind. Consciousness in this most general sense consists in one's possession of the ability to become *conscious of* aspects of a scene (that is, in one's ability to see, to explore aspects of the environment in a fashion mediated by the relevant sensorimotor contingencies).

6.3. The problem of qualia

As noted above in section 2.7, it may be argued that there is still something missing in the present account of vision, namely, an explanation of the *qualitative character* of visual experience. Can the sensorimotor contingency theory in addition provide an explanation of what philosophers have called "the raw feel" or "qualia" of seeing?

"Quale" is a technical term in philosophy. Like most such terms, to become clear about its precise meaning is to enter into the throes of philosophical controversy. Qualia are frequently characterized as the "phenomenal," or "qualitative," or "intrinsic" properties of experience, and they are typically contrasted with "intentional," or "representational," or "functional" features. Qualia are said to be that thanks to which there is something that it is like to have an experience (something that is, in addition, independent of representational or functional features). One of the central philosophical debates surrounding qualia concerns the question whether qualia can be studied by means of traditional biological and cognitive science. It has been suggested on this point that there is an unbridgeable "explanatory gap," that it is not possible to explain the subjective, felt aspects of experience in behavioral, physical, or functional terms.

In our view, the qualia debate rests on what Ryle (1949/1990) called a category mistake. Qualia are meant to be properties of experiential states or events. But experiences, we have argued, are not states. They are ways of acting. They are things we do. There is no introspectibly available property determining the character of one's experiential states, for there are no such states. Hence, there are, in this sense at least, no (visual) qualia. Qualia are an illusion, and the explanatory gap is no real gap at all.

It is important to stress that in saying this we are not denying that experience has a qualitative character. We have already said a good deal about the qualitative character of experience and how it is constituted by the character of the sensorimotor contingencies at play when we perceive. (We have more to say about this below; see also the discussion of the individuation of sensory modalities in sect. 2.5). Our claim, rather, is that it is confused to think of the qualitative character of experience in terms of the *occurrence* of something (whether in the mind or brain). Experience is something we do and its qualitative features are aspects of this activity.[41]

6.4. What gives rise to the illusion of qualia?

Many philosophers, vision scientists, and lay people will say that seeing always involves the occurrence of raw feels or qualia. If this view is mistaken, as we believe, then how can we explain its apparent plausibility to so many? In order to make our case convincing, we must address this question.

In our view, there are two main sources of the illusion. The first pertains to the unity and complexity of experience. We tend to overlook the complexity and heterogeneity of experience, and this makes it seem as if in experience there are *unified sensation-like* occurrences. The second source of illusion has to do with the *felt presence* of perceptible qualities. Because, when we see, we have continuous access to features of a scene, it is *as if* we continuously represent those features in consciousness. We elaborate these two mistaken lines of reasoning in turn.

6.4.1. The unity of experience. Scientists and philosophers frequently get the phenomenology of experience wrong; they misdescribe what perceptual experience is like. Consider, as an example, the experience of driving a Porsche and its distinctive qualitative character. What does this feeling consist of? What is it like? Notice that, in one sense, there is no feeling of Porsche-driving. That is, the character of Porsche-driving does not consist in the occurrence of a special sort of momentary flutter or bodily sensation.

What defines the character of driving a Porsche, rather, is something more complex. There are characteristic ways in which the vehicle accelerates in response to pressure on the gas pedal. There are definite features of the way the car handles turns, how smoothly one can change gears, and so on. What it is like to drive a Porsche is constituted by all these sensorimotor contingencies and by one's skillful mastery of them, – one's confident knowledge of how the car will respond to manipulations of its instruments.[42]

In one sense, then, there is no Porsche-driving quale. After all, what it is like to drive a Porsche depends on these various activities. In another sense, however, one can speak of the qualitative character of driving a Porsche, but this must be understood not in terms of the occurrence of a sensation-like quale in the mind, but rather, in terms of one's comfortable exercise of one's knowledge of the sensorimotor contingencies governing the behavior of the car.

We propose that the same account can be extended to such prototypical visual qualia as "the raw feel of a shade of red." Suppose you stand before a red wall. It fills up your field of view. What is it like for you to see this red wall? Try to describe the experience. How do you fulfill this instruction? One thing you might do is direct your attention to one aspect or another of the wall's redness. For example, you might focus on its hue, or its brightness. In this way you become transitively conscious of (that is to say, *aware of*) this or that aspect of the wall's color. How do you accomplish this? In what does your focusing on the red hue of the wall consist? It consists in the (implicit) knowledge associated with seeing redness: the knowledge that if you were to move your eyes, there would be changes in the incoming information that are typical of sampling with the eye; typical of the nonhomogeneous way the retina samples color; knowledge that if you were to move your eyes around, there might be changes in the incoming information typical of what happens when illumination is uneven, and so on. Importantly, there is not *one* thing in which the focussing of your attention on the hue (say) consists. Eye movements, shifts of attention, the application of understanding – seeing the red hue of the wall consists in all of this. There is no simple, unanalyzable core of the experience. There are just the different things we do when we see the redness of the wall.

In one sense, then, we can say that there is no red-quale (just as there is, in a sense, no Porsche-driving quale). An experience of a red surface is not a sensation-like occurrence. Seeing something red is a variegated activity, and to describe its character adequately, one must describe the many different things we do when we see something red.

6.4.2. The felt presence of qualities. Let us now turn to the second source of the illusion of qualia. Consider once again the phenomenon of change blindness. Many people say that they have the impression that when they see, the entire visual field is present to consciousness in all its nearly infinite detail. However, the change blindness results suggest that we do not have such detailed, picture-like awareness. What explains the conviction that we do? As we have discussed above, and as argued by O'Regan (1992; cf. also O'Regan et al. 1999), the explanation is that we have access to all the detail by means of the mere flick of an eye or turn of the head, and so it is *as if* we had everything in view all the time. The *feeling* of the presence of detail stems from our implicit knowledge of the ways in which movements of the eye and head gives rise to new detail and new information. Importantly, one can explain this feeling without supposing that all the detail is represented in consciousness.

In exactly this way, when we see something red we *feel* that the redness has a certain definite, sensation-like presence and immediacy. The explanation for this is that we have access to the redness by the most minute of eye movements or attentional shifts. The redness is there, in the environment. The slightest eye, head, or attention movement reveals further information about its character. Because we have continuous *access* to the redness in the environment, it may seem as if we are mentally in contact with it continuously. This leads us to say, mistakenly, that there is a feeling of redness (say) in our heads all along.

6.5. Is the illusion of qualia really so widespread?

Is the illusion of qualia really as widespread as it would seem? Perhaps not. If you ask what a person sees, he or she will not bring up visual experiences and their intrinsic features. In everyday life, discussions of what we see are for the most part confined to discussions of things themselves (of the things we see). Even when we are viewing a piece of art, when we may deliberately try to reflect on the way the work affects us visually, nonphilosophers will rarely confuse the question *what it is like to look at the piece* (what it reminds one of, how it makes one feel, whether one finds it pleasant, or not) with that favorite question of philosophers, namely, *what is it like to have an experience as of seeing a painting* (that is, what are the intrinsic, qualitative features of the visual experience)?

Another way to put this point is to say that qualia-based accounts of the phenomenology of experience actually misdescribe the phenomenological character of experience (what experience is like). Qualia-talk, one might say, is theory driven and the illusion of qualia is a theoretical illusion. Crucially, normal perceivers do not, by virtue of being normal perceivers, buy into the relevant theory.

6.6. The ineffability of the qualitative character of experience

We have proposed that experience is a temporally extended activity of exploration mediated by the perceiver's knowledge of sensorimotor contingencies. The differences in the qualitative character of perceptual experiences correspond to differences in the character of the relevant sensorimotor contingencies. Just as the difference between driving a Porsche and driving a tank consists in the different things you do in driving it – that is, in the different skill-based understanding of how to drive the vehicle, – so the difference between seeing a red flower and smelling a red flower consists in the different patterns of sensorimotor contingency governing one's perceptual encounter in each situation. To experience a red object, or the feel of driving a Porsche, is to know, for example, that if you change the illumination in such and such ways (or press down on the accelerator in such and such ways), it will produce such and such changes in the stimulation.

It follows, according to this view, that to reflect on the character of one's experience is to reflect on the character of one's law-governed exploration of the environment, on what one does in seeing. Some of the sensorimotor contingencies governing vision are easily accessible to awareness. If you reflect on the character of your visual experience of a colorful flower, for example, it is easy to comprehend the manner in which the appearance of the flower is a function of viewing angle and illumination. If you look at a plate and turn it, you can become aware of the way its profile becomes elliptical. If you put on inverting lenses, it is immediately apparent that eye and head movements produce surprising patterns, thus enabling us to direct our attention to the disruption of familiar patterns of sensorimotor contingency. But though we have access to these aspects of the sensorimotor contingencies, there are other components of the sensorimotor contingencies which do not lend themselves easily to propositional description, and which are not so easily brought into consciousness: the exact laws that the flower's color obeys when you change the illumination, the exact rule determining the modification of the plate's profile, the precise disruption caused by distorting lenses. Other examples even less accessible to consciousness are: the particular way the macular pigment and the non-homogeneity of retinal sampling affect sensory input when the eye moves; the optic flow that occurs when the head rotates, and so on.

We believe that these considerations enable us to get clear about a feature of experience that has often provoked puzzlement on the part of scientists and philosophers, namely, its apparent ineffability. It is very difficult to describe everything we do when we see, just as it is difficult to describe everything we do when we are engaged in other skillful activities such as athletic endeavors, playing an instrument, or speaking a language. A major portion of our mastery of sensorimotor contingencies takes the form of practical know-how. When we attempt to inquire into the more subtle features of what goes on when we perceive, we immediately come up against the fact that it is very difficult to describe any but the most high-level, gross sensorimotor contingencies.

There is nothing mysterious about this inability. In general, the ability to know how to do something does not carry with it the ability to reflect on what it is one does when exercising the ability in question. The difficulty in describing the character of experience is not evidence of the special character of experience in the world order. But it does bring forcibly to mind the fact that experiences are "exercisings" of complicated capacities, not ongoing occurrences in the mind or brain.

6.7. On the possibility of phenomenology

We hope it is clear that it is no part of our argument to deny the possibility of, or the importance of, phenomenological reflection on experience.[43] Indeed, we believe that our view provides an account of the subject matter of phenomenology that is superior to that put forward by qualia-oriented positions.

First, our theory is supported by careful reflection on what it is like to have perceptual experience. It is commonly asserted by both philosophers and scientists, that it seems to normal perceivers as if perception involves detailed internal representations of the environment in the head. As noted in section 4.2, we believe this misdescribes the character of seeing. First of all, in seeing we commit ourselves to no beliefs about what is going on in our heads. Seeing is directed to the world, not the brain. Second, when we see, we take the perceived detail to be out there in the world, not in our head. Indeed, we take ourselves to be embedded in the environment and to have access to detail through active exploration. In our view, it is just bad phenomenology to assert that we take ourselves to have a 3D-model or picture in the head when we see. In short, we believe that, once it has broken free of clichés about pictures in the head, phenomenological reflection on the character of experience does support the kind of approach developed here.

Second, traditional qualia-based approaches to experience threaten to make experience itself something mysterious and inaccessible. However, the subject matter of phenomenological reflection is not an ephemeral, ineffable, sensation-like momentary occurrence in the mind, but, rather, the real-world, temporally extended activity of exploring the environment and the structure of sensorimotor contingencies. There is a qualitative or phenomenological difference between seeing and hearing and touching, as stated. These are different activities, corresponding to different modes of exploration of the structure of sensory-motor contingencies. To see a bottle, for example, is to explore visual-motor contingencies, such as transformations in the appearance of the bottle as one moves in relation to it. To touch it, on the other hand, is to explore the structure of tactile-motor contingencies. The bottle impedes, guides, and informs tactile exploration of the bottle. To reflect, then, on what it is like to see the bottle, or to touch it, is to reflect on just these sorts of facts about the active engagement the perceiver undertakes with the environment (see Noë 2001). In this way, we believe that the kind of approach we lay out in this paper helps place phenomenology as an undertaking on solid ground (see Noë 2002, for a development of this idea).

6.8. Overcoming the explanatory gap (or, Why there is no gap)

As noted above, the problem of the explanatory gap is that of explaining qualia in physical or biological terms. We believe that our view bridges this gap. More accurately, it demonstrates that the gap itself is an artifact of a certain – we believe mistaken – conception of experience. There is not really any gap at all.

Our claim, simply put, is this: there is no explanatory gap because there is nothing answering to the theorist's notion of qualia. That is, we reject the conception of experience that is presupposed by the problem of the explanatory gap. (Note that we can make this claim even though we do not deny, as we have been at pains to explain above, that there are experiences and that experience has qualitative character.)

To appreciate the structure of our claim, consider once again, very briefly, the Porsche-driving example. We have argued that the feeling of driving a Porsche derives from the different things we do when we drive a Porsche, and from our confident mastery of the relevant sensorimotor contingencies. We can now appreciate that there is no need to explain the physical or causal basis of the occurrence of the unitary Porsche-driving quality, for there is no such quality. What does need to be explained is the physical (neural) basis of the various component skills that are drawn into play when one drives a Porsche (for it is these that constitute the feeling). And so, likewise, there is no need to seek a neural basis for the occurrence of visual qualia such as that of red, for, in the relevant sense, there are no such qualia.

To this it will be objected that it is no more easy to see how possession and mastery of sensorimotor skills is to bridge the explanatory gap, than it is to see how different patterns of neural activity can accomplish the same feat. But this very question betrays a failure to understand our proposal. For our claim is *not* that knowledge and exercise of sensorimotor contingencies can solve the same feat. Our claim is that there is no feat to be accomplished and, *therefore,* no possible way in which neural activity can accomplish it. Let's return again to simple examples. You hold a bottle in your hand. You *feel* the whole bottle. But you only make contact with isolated parts of its surface with isolated parts of the surface of your hands. But don't you feel the whole bottle as present? That is, phenomenologically speaking, the feeling of presence of the bottle is not a *conjecture* or an *inference.* The feeling you have is the knowledge that movements of the hand open up and reveal new aspects of bottle surface. It feels to you as if there's stuff there to be touched by movement of the hands. That's what the feeling of the presence of the bottle consists in. But the basis of the feeling, then, is not something occurring now. The basis rather is one's knowledge *now* as to what one can do.

6.9. Summary

Let us summarize the main claims of this section.

(1) There are two kinds of visual consciousness. There is transitive visual consciousness (or *consciousness of*), which consists in one's awareness of an aspect of a scene. There is *visual consciousness in general,* which consists in one's general capacity to become aware of different features of the scene. Transitive consciousness, as a form of awareness, can be explained just as we explain visual awareness in section 2.6 above. To be aware of a feature is to exercise one's practical knowledge of the relevant sensorimotor contingencies. Visual consciousness in general is just the higher-order capacity to exercise such mastery.

(2) The difference between different perceptual experiences, and between different perceptual experiences in different sensory modalities, can be explained in terms of the different things we do in having the experience and in terms of the different rules of sensorimotor contingency that are invoked in each case. The supposition that there are further qualitative aspects of experience that cannot be explained along such lines is an illusion, engendered by: (a) our tendency to fail to attend to the heterogeneity and complexity of experience; (b) our tendency to treat continuous access to environmental detail as the continuous representation of that detail. Moreover, we claim that the illusion of qualia is actually not as widespread as philosophers often suggest, and that the conception of experience we develop in this paper – experience as a mode of skillful activity – is actually truer to the actual character of felt experience than qualia based views.

(3) There is no explanatory gap. We do *not* claim that it is possible to explain the physical basis of conscious experience by appeal to sensorimotor contingencies. How, one might ask, can sensorimotor contingencies explain phenomenal consciousness any better than other proposals that have been made? Rather, we argue, as should by now be clear, that the conception of phenomenal consciousness itself must be (and can be) rejected, and so there is no longer any puzzle about how to explain that. As we make clear in the points above, other aspects of consciousness can indeed be explained according to our view.

We have not attempted to present solutions to such philosophical chestnuts as the problem of undetectable spectrum inversion, or the problem of zombies. Instead, we have turned our attention to the presentation of a framework within which to investigate the nature of vision and visual consciousness. We have drawn attention to the wealth of empirical data that support our theory. In addition, we have tried to provide some statement of what we take to be the implications of this view for progress on the topic of visual consciousness. We have adopted the strategy of trying to demonstrate the fruitfulness of our approach instead of that of refuting the philosophical opposition. This sort of indirect approach is necessary when what divides camps is not so much disputes over what the facts are, but rather, fairly messy questions about how to make sense of the interdependence of a whole network of related ideas: seeing, vision, visual experience, visual consciousness, qualia, raw feel, awareness, and attention. We have made a number of proposals about how to think about this raft of interconnected phenomena which will, we hope, allow for empirical progress.

7. Philosophical niceties

7.1. Awareness versus consciousness

Chalmers (1996a) distinguishes between awareness and consciousness. Awareness, according to Chalmers, is a state in which some information (that of which we are aware) is available for control of behavior and for guiding verbal report. Consciousness, or *experience,* on the other hand, is an intrinsically qualitative state whose links to behavior are inessential. Chalmers' distinction is very similar to Block's (1995b) distinction between access-consciousness and phenomenal-consciousness. A state is access-conscious, according to Block, if it is poised to be used to govern rational thought, guide behavior, or give rise to verbal report. A state is phenomenally conscious, however, if it *is an experience.* Block and Chalmers agree that awareness (or A-consciousness) is a functional notion, definable in terms of behavior and dispositions to behave, and they agree that consciousness (or P-consciousness, or experience) are non-functional notions (that is, functional duplicates can differ in their P-consciousness). (Block and Chalmers differ in important details that do not concern us here.)

We are skeptics about phenomenal consciousness understood the way Block and Chalmers understand it. As we stated above, what explains the illusion that seeing consists in the occurrence of an *internal qualitative state* is the fact that, at any moment, one can direct one's attention to one's activity of looking and so encounter such qualities as the redness of a wall, or the distinctive shape of a seen object. Moreover, we are able to track not only objects of awareness, but our tracking activity itself and thus become aware (in the functional sense) of the percepts induced by the patterns of sensorimotor contingency governing our seeing. The experience of red, for example, arises when we know (though this is not propositional, but rather, practical knowledge) that, for example, if we move our eyes over a red region, there will occur changes typical of what happens when our non-homogeneously sampling retinas move over things whose color is red. It is, then, our continuous *access* to the redness that provides the key to understanding why it (mistakenly) seems to us as if we are continuously undergoing experience as of something red.

Our account of seeing and visual awareness thus cuts across the distinction between awareness and consciousness (as Chalmers puts it), or between A- and P-consciousness (in Block's terms). Visual experience is a matter of access, but access to the world, and to one's activity of tracking and interacting with the surrounding scene, not to one's internal information-bearing states. The felt or qualitative character of seeing is to be explained in terms of this active conception.

7.2. Blindsight

Block (1995b) puts the concepts of A- and P-consciousness to work in his discussion of blindsight. Patients with blindsight have suffered lesions in the visual cortex as a consequence of which they appear to be blind in a region of the visual field (Pöppel et al. 1973; Weiskrantz 1986). Subjects report that they see nothing when a stimulus is presented to their scotoma. When asked to guess (from a number of choices) what is present in their blind field, subjects are correct at a rate well above chance. There would seem to be a sense, then, in which these individuals see without seeing. One possibility is that these patients see, but are unconscious of seeing. This in turn suggests that the function of consciousness is to enable us to make use of the information we acquire. Block (1995b) attacks this reasoning because it fails to distinguish between A- and P-consciousness. P-consciousness is surely lacking, but A-consciousness is absent too, at least for the most part. One is not entitled, then, to draw general conclusions about consciousness from the phenomenon of blindsight.

Block contrasts blindsight with a nonactual but, he thinks, conceptually possible phenomenon of superblindsight. In superblindsight, as Block describes it, subjects have apparently normal access to information acquired in their blind fields, but they lack *experience* of the information. He invites us to imagine that these individuals have been trained to trust their "guesses" about what is present in their blind fields. This information, therefore, is available to guide action and speech. Indeed, we are asked to imagine that, as far as speech and behavior go, people with superblindsight seem normal. There is one noteworthy exception, of course. If you ask them whether they visually experience what is present in their blind field, the way they experience what is present to their non-blind field, they reply that they experience nothing. They are as good as blind as far as feeling goes.

Block's main contention then is two-fold: (1) that superblindsight is visual A-consciousness in the absence of P-consciousness; and (2) that superblindsight is conceptually possible. We doubt both points. As for (1), it seems that we have grounds for doubting that the patient really has a blind field. After all, the patient appears to see just fine. As Dennett (1995) notes, Block's account appears to trade, illegitimately, on the fact that in actual blindsight the kinds of information involved are remarkably sparse (on this point, see also Noë 1997). The subject is correct, for example, about the orientation of a line grating. But if we imagine informational content to be greatly enriched, as would seem required in the case of superblindsight, then the claim, on the part of the subject, that he lacks P-consciousness, becomes highly implausible. It is difficult to make sense of the claim that a person might offer an accurate description of a painting, say, describing all the colors and the geometry of the composition in a natural manner, all the while having no experience of the painting. One loses all grip on what it could mean to say that the subject *has no experience.* And this indicates the nature of our misgivings about (2). If you are perceptually alert to the presence of environmental detail in a manner that allows you to describe what is present, and if you are sensitive to the appropriate visual laws of sensorimotor contingency (for example, if the detail is no longer accessible when you close your eyes or the lights go out, if the image shifts in the normal way when you move your eyes, if your attention is immediately drawn to any change in the image, etc.), – then surely it would be very peculiar to say that you are not experiencing/seeing the painting (Noë 1997).

Nor do we find Block's (1995b) examples of cases of P-consciousness without A-consciousness convincing. He gives the example of having a conversation while a power drill makes a racket outside the window. One is engrossed in the conversation and one does not notice the drill. All of a sudden one notices it. Block proposes that in a case such as this, insofar as one did hear the drill before noticing it, one was P-conscious of the drill while at the same time A-unconscious of it. When one noticed the drill, one becomes A-conscious of what one had previously heard and been P-conscious of all the time.

But did one hear the drill before one noticed it? The view developed in our paper here requires a negative answer to this question. One does not hear the sound of the drill because one does not make use of one's auditory tracking. This is of course compatible with its being the case that we are sensitive to the sound before we hear it (before we become conscious of it). The auditory system will analyze and store (perhaps only in a short-term memory buffer) information pertaining to the drill. But we do not use that information, nor are we, before we notice the drill, *poised* to use that information or *able* to use that information to guide our behavior, thought, movement, or perceptual exploration.

One might challenge Block's view in another way as well. Consider a slightly different but familiar example. A bell is chiming. All of a sudden you notice not only that there is a bell chiming, but that there were six chimes in all. Surely this shows that you heard the chimes even before you noticed it? Indeed, what this would show, as Chalmers has argued, is that there is a sense in which one *was* in fact *poised* to make use of the unnoticed sounds one was hearing even as one failed to notice them. That is, according to this line of reasoning, one was A-conscious of the unexperienced sounds (contrary to what Block would say).

One virtue of this account is that it perhaps fits somewhat better with ordinary usage of words like "hear" and "see." That is, it seems quite natural to say that you heard the clock chime without noticing it. But there are substantive empirical reasons to reject this account nonetheless. The fact that a stimulus is present and is actively impinging on the senses, does not entail that you perceive it. This is the central upshot of the change blindness studies (discussed in sect. 5.10) and also recent work on so-called "inattentional blindness" (discussed in sect. 5.11). The fact that a stimulus is present means that it is available to be probed by the active animal. Only while the active probe is occurring do you get conscious perception (seeing or hearing, say).

The conflict between our view and that of common sense is actually more apparent than real. As we noted earlier in our discussion of awareness, awareness is a matter of degree. Part of what makes it seem so reasonable to say that you heard the noise without noticing it, or that you (a driver) saw a car without noticing it, is that we may call to mind cases where you are in fact noticing a sound or an object *a little bit.* For example, you are trying to have a conversation and there's that irritating noise in the background which threatens to interrupt you but to which you are paying very little attention. Nevertheless, having said all this, we are quite prepared to bite the bullet and insist that in the complete absence of *current access,* there is no perception.

Note that to say that there is no perception is not to say that there may not be significant *unconscious* influence on behavior or action.[44]

7.3. Our relationship to Dennett

The view developed in this paper is very similar in important respects to the position developed over the last few decades by the philosopher D. C. Dennett (Dennett 1978; 1987; 1991; Dennett & Kinsbourne 1992). But, as the discussion of the previous section suggests, there are important differences as well.

Many philosophers and scientists assume that consciousness is an *intrinsic* property of neural states. The idea is that, among the multitude of content-bearing states in the brain, some subset of states have an additional property of being *phenomenally* conscious. (This is in contrast to states which, in the terminology of Block, are *access*-conscious. This access consciousness is not thought to be an intrinsic property of the state but one that depends on the relation between that state and others in the broader system.) The problem of consciousness, in this general picture, is to understand what processes or mechanisms or events in the brain make certain contents phenomenally conscious. Where, and how, does consciousness happen in the brain?

We reject not only specific attempts to answer this question (oscillations, synchrony, microtubules, etc.), but the assumptions implicit in the question itself. That is, like Dennett, we reject Cartesian materialism (Dennett 1991; Dennett & Kinsbourne 1992). Phenomenal consciousness is not a property of states in what Dennett calls the subpersonal system (i.e., the brain – whether thought of in neural, or in more abstract, cognitive or computational terms) (Dennett 1978; 1969; 1987). There need be no one-to-one correlation between states of consciousness and events in the brain.

But this brings us to the main point of our disagreement with Dennett. Although we reject accounts of phenomenal consciousness as a property of subpersonal states, we do not deny (as we have sought to make clear in the previous sections, especially sect. 6.7), that there are experiences and that there are facts about what experiences are like. But these, however, are facts not about a person's qualia or raw feels. They pertain, rather, to the person's (or animal's) active engagement with the world he or it inhabits.[45] They are facts at the personal (as opposed to subpersonal) level. We return to this point below.

One of the cornerstone's of Dennett's approach to the problem of consciousness, is his conception of *heterophenomenology* (Dennett 1991). In many respects, we are very sympathetic to this approach. The best way to understand what Dennett means by heterophenomenology is to contrast this view with could be called *introspectionism.* Introspectionism is the view that the conscious subject has immediate epistemic access to his or her conscious states. Perhaps not too many writers would endorse introspectionism when put forward in this blunt manner, but it is clear that something like this idea drives a good deal of discussion in contemporary consciousness studies. Theorists believe that we know, on the basis of reflection on our own case, what our own conscious states are like. Dennett rejects introspectionism. Dennett has a lot to say about why introspectionism is untenable, and we are sympathetic to his position. For our purposes it is enough to point out that, according to Dennett, as scientists we cannot assume that subjects are right in their first-personal avowals of conscious experience. Such reports are just further bits of evidence about the nature of mental life and they have no privileged status with respect to other forms of evidence (e.g., psychophysical, neural, psychological, etc.). According to heterophenomenology, then, first-person reports of experience have no special status attached to them. There is no deep and unfathomable asymmetry between what can be known *in the first person,* and what can be known *in the third person.*

Although we endorse Dennett's rejection of naïve introspectionism, our endorsement of the claim that first-person approaches to consciousness are not privileged with respect to third-person approaches is guarded. To appreciate why, consider an example: Dennett (1991) criticizes what he takes to be the widespread assumption, on the part of perceivers, that the visual field is in sharp detail and uniform focus from the center out to the periphery. Simple tests (e.g., the colored pencil test mentioned earlier in sect. 5.6), and well-known facts about the non-homogeneity of the retina, suffice to show that this account of the quality of the visual field is misguided. But is it really true that normal perceivers think of their visual fields this way? Do normal perceivers really make this error? We think not. As noted earlier in connection with change blindness (and see Noë 2001; Noë et al. 2000; and Pessoa et al. 1998), normal perceivers do not have ideological commitments concerning the resolution of the visual field. Rather, they take the world to be solid, dense, detailed, and present, and they take themselves to be embedded in and thus to have access to the world.

The point of this example is that Dennett seems to *mis*-characterize how things seem to perceivers, that is, he mischaracterizes their first-person judgments as to the quality of experience. He does this precisely because he is insufficiently attentive to the actual phenomenology of experience. What this shows is that there are substantive empirical questions about the first-person quality of experience. To investigate such questions, presumably, one must avail oneself of the first-person perspective. From this it does not follow, to be sure, that first-person methods are privileged with respect to third-person methods, but it does follow that it ought to be possible to develop modes of first-person investigation of experience that do not suffer from the flaws of qualia-based (introspectionist) approaches.[46]

The crucial point is that nothing in Dennett's criticisms of naïve introspectionism entails that all first-person approaches to consciousness must take the form of naïve introspectionism, and so nothing in the arguments speaks against the possibility of, or importance of, first person approaches. In the concept of perceptual experience we have endorsed, first-person reflection on the character of experience would not consist in *intro*spection at all, but rather in attentiveness to the complexity of the activity of perceptual exploration. Ironically, as we have seen, Dennett's rejection of the importance of the first-person perspective has led him, at crucial junctures, to *mis*describe the character of perceptual experience.

8. Visual neuroscience

8.1. The brain and vision

Much work on visual neuroscience rests on the idea that for every perceptual state there is a neural correlate sufficient to produce it. In addition, it is widely supposed that the function of this neural substrate is to produce sensory experience by generating a "representation" corresponding to the content of the experience. A very different conception of the role of the brain in vision emerges from the standpoint of the sensorimotor contingency theory.

According to this theory, seeing is a skillful activity whereby one explores the world, drawing on one's mastery of the relevant laws of sensorimotor contingency. Seeing, in this sense, is somewhat like dancing with a partner. Dancing is a complicated activity. There is no *one* thing in which dancing consists and there is no single state of being in the dancing state. Dancing consists in the integration of a range of connected skills: sensitive listening, coordinated movement (or sometimes the *absence* of movement); and, importantly, partner dancing requires the presence of a partner to whose actions and reactions one is appropriately attuned. There is no doubt that neural activity in the brain is necessary to enable one's skillful performance of the dance. But this neural activity is not *sufficient* to produce the dancing. This is so because the accompanying, appropriate actions and reactions of the partner are also needed. These provoke weight changes, disequilibria, rebounds, and so on, which cannot occur without the partner being present, and which are part and parcel of the dancing activity.

In the same way, we argue, seeing also necessarily involves particular forms of action and reaction on the part of the visual apparatus and the environment. The brain enables us to see by subserving the different capacities that get drawn on in the activity of visual exploration. But the brain's activation does not in itself constitute the seeing. In partner dancing, specifying the bodily configuration or brain state of the dancer is not sufficient to specify the dance (because we need additionally to know how the partner is currently interacting). Likewise, in seeing, specifying the brain state is not sufficient to determine the sensory experience, because we need to know how the visual apparatus and the environment are currently interacting. There can therefore be no one-to-one correspondence between visual experience and neural activations. Seeing is not constituted by activation of neural representations. Exactly the same neural state can underlie different experiences, just as the same body position can be part of different dances.

How then are we to understand the role played by the brain in vision? Our proposal, which we develop below, is that the brain supports vision by enabling mastery and exercise of knowledge of sensorimotor contingencies.

8.2. The search for neural representations and the neural correlate of consciousness

Perhaps the most widely cited work which might be thought to constitute evidence for the existence of cortical representations of sensory stimuli could be taken to be the observations of Penfield, who solicited sensory responses from unanesthetized patients undergoing brain stimulation (e.g., Penfield & Jasper 1954). More recent work in visual science and consciousness studies has been devoted to the quest for what has been called "neural correlates of consciousness" (Crick & Koch 1990; 1995; 1998; – for an illuminating review, cf. Chalmers 1996a, Ch. 6). As an illustration of such work, we can use the impressive studies of Logothetis and colleagues (Leopold & Logothetis 1996; Logothetis 1998; Logothetis et al. 1996) analyzing neural substrates of binocular rivalry in laboratory monkeys. In binocular rivalry, each eye is presented with a different stimulus (e.g., a horizontal bar, a face). Under these conditions the observer experiences not both stimuli, or some amalgam of the two, but rather a sequence of alternating percepts corresponding to one or other of the two stimuli. When one stimulus is dominant, the other is not perceived. The perceptual reversals occur irregularly and at intervals of a few seconds. Logothetis and collaborators show that in tested visual areas (e.g., V1/V2, V4, MT, IT, STS), some neurons are unaffected by perceptual reversals. The activity of these neurons is driven by the stimulus patterns entering the eyes, which remain unchanged. The activity of other neurons, however, depends directly on the internally generated shifts in the percept. The percentage of such percept-driven cells is substantially higher in IT and STS – where 90% of tested neurons correlate to percepts, – than in other visual areas. (In V1/V2, for example, a much smaller percentage of neurons were percept-driven.) These data suggest (it is claimed) that neural activity in IT and STS forms the neural correlate of the experience.

Other kinds of neural representations or neural correlates of conscious perceptual experience arise in the context of perceptual completion phenomena. A classical example is the work of von der Heydt and his colleagues, who found neurons in V2 that fire for illusory contours in a very similar way that they fire for real contours (Peterhans & von der Heydt 1989; von der Heydt & Peterhans 1989; von der Heydt et al. 1984). A number of other examples involving perceptual completion have been reviewed by Pessoa et al. (1998).

Work like that described above has been received with enthusiasm: researchers believe that the discovery of neural representations that correlate with perceptual experience brings us closer to understanding what gives rise to the perceptual experience. The underlying assumption is that if a set of neurons is found in the brain which correlates strongly with aware perceptual states, then, because these neurons are probably linked to the mechanisms that are generating awareness, we are likely to be able to explain perceptual awareness by appeal to this neural activity.

But this reasoning is unsound. Indeed, consider what would happen if we were to actually find a set of neurons that correlated *perfectly* with visual awareness. For the sake of illustration, suppose we were to discover that in the pineal gland of macaque monkeys there was a tiny projection room in which what is seen by the monkey was projected onto an internal screen whose activity correlated perfectly with the monkey's visual awareness. On reflection it is clear that such a discovery (which would surely be the Holy Grail of a neural correlate of consciousness seeker!) would not bring us any closer to understanding how monkeys see. For we would still lack an explanation of how the image in the pineal gland *generates* seeing; that is, how it enables or controls or modulates the forms of activity in which seeing consists. We would certainly be entitled, on the basis of the strong correlation between features of what is seen and features of what is projected onto the pineal projection screen, to assume that this neural activity played some role in vision. But nothing more could be said about such a discovery.

Why do some researchers believe that to understand the nature of consciousness or vision it is necessary to track down the neural representations that correlate with conscious experience? One possible explanation is that these researchers are (perhaps unwittingly) committed to the idea that the discovery of *perfect* correlation would give us reason to believe that we had discovered *the* neural activity sufficient to produce the experience (as suggested by Chalmers 1996a). Teller and Pugh (1983) call such a neural substrate of experience the bridge locus. In addition, thinkers may unwittingly subscribe to what Pessoa et al. (1998) have called *analytic isomorphism.* This is the view that for every experience there will be a neural substrate whose activity is sufficient to produce that experience (a bridge locus), and that there will be an isomorphism (though not necessarily spatial or topographic) between features of the experience and features of the bridge locus. It is the existence of such an *isomorphism* that works to justify the claim that the discovery of such a neural substrate would *explain* the occurrence of the percept.

We believe that one must reject the metaphysical dogma of analytic isomorphism. As argued by Pessoa et al., no neural state will be sufficient to produce experience. Just as mechanical activity in the engine of a car is not sufficient to guarantee driving activity (suppose the car is in a swamp, or suspended by a magnet), so neural activity alone is not sufficient to produce vision.

Note also that if this view is correct, then it is a mistake to expect to find neurons which *are* perfectly correlated with visual consciousness. Ultimately, visual consciousness is not a single thing, but rather a collection of task and environment-contingent capacities, each of which can be appropriately deployed when necessary. Furthermore, we expect that if neurophysiologists do find neurons that correlate strongly with awareness, then most likely this will only be for one or another set of conditions or tasks.

8.3. There is no need for "binding"

Neuroanatomists believe that the visual system is composed of numerous, more or less independent subsystems (or modules), which extract a variety of different attributes such as color, contrast, depth, orientation, and texture from the visual stimulus (e.g., De Yoe & van Essen 1988; Livingstone & Hubel 1988; Zeki 1993). The fact that these modules operate independently and are often localized in different cerebral regions, raises the question of how the separate streams of information ultimately come together to give us the unified perception of reality that we subjectively experience. One suggestion for solving this so-called "binding problem" was the idea of the "grandmother cell" in which single cells, or at least highly localized cerebral regions, combine information pertaining to specific percepts: for example, face-sensitive cells (Rolls 1992); place sensitive cells (O'Keefe et al. 1998); view sensitive cells (Rolls & O'Mara 1995). A more recent idea which does not require bringing signals into a single brain location has also received support from neurophysiological evidence (cf. Abeles & Prut 1996; Brecht et al. 1998; Castelo-Branco et al. 1998; Gray & Singer 1989; Llinas & Ribary 1993). According to this view, separate cortical areas which are concurrently analyzing the different aspects of a stimulus might oscillate in synchrony, and it might be this synchrony which provides the perceptual experience of unity.

There are two motivations in the reasoning which underlies these types of investigations: one concerns temporal unity, and the other concerns "conceptual" unity.

Certainly it is true that when we recognize an object, we have the impression that all its attributes are seen simultaneously at one "perceptual moment." This leads scientists to think that the objects' attributes must be bound together *synchronously* in the internal representation in order to provide the singleness of the perceptual moment. But this is a fallacy. Thinking that physical synchrony is necessary for having a synchronous experience is the same kind of fallacy as thinking that because things look like 3D models or picture postcards to us, there must be a topologically equivalent map in the brain. Underlying this fallacy is the implicit assumption that the synchrony or coherence of perception requires presenting information in a synchronous or coherent way to an internal homunculus. In fact, just as the perception of the 3D world does not require 3D maps in the brain, subjective simultaneity does not require simultaneity of brain events.[47] This point has been made by Dennett and Kinsbourne (1992; see also O'Regan 1992; Pessoa et al. 1998). What explains the temporal unity of experience is the fact that experience is a thing we are doing, and we are doing it *now.*

Coming now to the issue of "conceptual" coherence, a similar argument can be made: the fact that object attributes seem perceptually to be part of a single object does not require them to be "represented" in any unified kind of way, for example, at a single location in the brain, or by a single process. They *may* be so represented, but there is no *logical necessity* for this. Furthermore, if they are represented in a spatially or temporally localized way, *the fact that they are so represented cannot in itself* be what explains the spatial, temporal or "conceptual" phenomenology of perceptual coherence.[48] What explains the conceptual unity of experience is the fact that experience is a thing we are doing, and we are doing it with respect to a conceptually unified external object.

We noted above that were we to discover pictures in the brain that correlated with the experience of seeing, we would still not have moved much closer towards an explanation of seeing. But once we recognize this, then we further realize that there is no reason to suppose that to explain seeing we should seek for detailed internal pictures. There is no longer any rationale for supposing that there is a place in the brain where different streams of information are brought together and "unified" (whether conceptually or temporally). With the appreciation of this point we can dismiss the problem of binding as, in essence, a pseudoproblem.

8.4. A new way of thinking about the role of the brain in vision: A program for future research

We have already taken steps toward a positive characterization of the role of the brain in vision by claiming (as we have in sect. 8.1 above) that studies of the neural bases of vision must be framed by a consideration of the whole animal's broader behavioral and cognitive capacities. In this section we try to extend these remarks.

Consider the missile guidance system we discussed in section 2.4. Suppose that at the present moment the target airplane happens to have gone out of the field of view of the missile. No information, let us suppose, is coming into the missile's sights right now. Nevertheless, the missile guidance system has a certain potential: it "knows" that by making the appropriate change in its trajectory, it should be able to bring the missile back into view. Thus, even though at this particular moment the airplane is not visible and no visual information is coming in, it is still correct to say that the missile is currently tracking its target.

Exactly the same point, we argue, can be made about seeing and the sensorimotor contingencies governing seeing. When you make an eye saccade, the sensory stimulation provided by an object will change drastically due to very strong retinal smearing. At that very moment you do not receive sensory input from the object. But there is no more reason to think that this interruption in stimulation leads to an interruption in seeing, than there is to think that the missile is no longer tracking the plane when the plane happens to go out of the missile's sights.[49] The missile continues to track the plane, and the perceiver continues to see, because each is master of the relevant sensorimotor contingencies and each is exercising those capacities in an appropriate manner. Seeing an object consists precisely in the knowledge of the relevant sensorimotor contingencies – that is, in being able to exercise one's mastery of the fact that if, among other things, you make an eye movement, the stimulus will change in the particular way typical of what happens when you move your eyes. If the stimulation due to the object did not change in that way, then you would not be seeing the object – you might, for example, be hallucinating it.

These considerations call attention to the fact that interruptions and discontinuities in stimulation (owing to saccades, blinks, eye movements, chromatic aberrations, and other supposed defects of the visual apparatus) are in fact part of what seeing is. It is one's exercise of the mastery of just such regularities in sensorimotor contingencies in which seeing consists. What is striking for present purposes is that just as moments of stillness and inactivity may be essential to the performance of a dance, so moments of neural inactivity may be precisely what characterizes the exercise of sight. This is a fact that can only come into focus through a conception of vision as a mode of activity such as that developed by the sensorimotor contingency theory.

Considerations such as these show further, that although neural activity is necessary for vision, there need be no one-to-one mapping between seeing and occurrent neural states and processes. Vision requires all manner of neural events, but crucially, in our view, the experience of seeing itself cannot be equated with the simultaneous occurrence of any particular neural activity. This follows from the fact that, at any given moment, the brain may be inactive.

What then is the function of the brain in vision? Very generally speaking, it is to enable the knowledge and exercise of sensorimotor contingencies. Seeing, we argue, is constituted by the brain's present attunement to the changes that would occur as a consequence of an action on the part of the perceiver. Visual experience is just the exercise of the mastery of relevant sensorimotor contingencies. An example may help to make the point clearer. Your visual apprehension of the roundness of a plate consists in part in your knowledge that changes in your relation to the plate (movements relative to the plate) will induce changes in the plate's profile. That it looks round to you now, despite its elliptical profile, is constituted by your application, now, of skillful mastery of the appropriate rule of sensorimotor contingency. Other rules of sensorimotor contingency may be, as it were, more low level. As you move your eye across a straight line (as discussed in sect. 2.2), there is a characteristic pattern of transformation of the retinal stimulation. The brain is attuned to this pattern. One important function of the brain may thus consist in the testing of the appropriateness of the application of certain patterns of sensorimotor contingency.

An important advantage of this view is that it allows us to escape from the problem of having to explain how brain activity could give rise to experience. We escape from this problem because we propose that experience does *not* derive from brain activity. Experience is just the activity in which the exploring of the environment consists. The experience lies in the doing.

8.5. Toward a sensorimotor approach to visual neuroscience

A good deal of recent neuroscientific research shows that to understand the role the brain plays in supporting perceptual and motor capacities, it is necessary to keep clearly in view the broader context of the animal's skillful task-oriented activity. Specific neural states cannot be perfectly correlated with specific perceptual states. You cannot understand the contribution of neural activity if you restrict yourself to a brain's-eye view. This fits with our model of vision and visual consciousness. Seeing is not a simple occurrence; it is a rich, exploratory activity within a certain environment and with certain sensory apparatus, drawing on a number of heterogeneous capacities. Neural activity does not in itself produce experience. Neural activity contributes to experience only as enabling mastery and exercise of the laws of sensorimotor contingency.

An exhaustive survey of this neuroscientific research goes beyond the scope of the present discussion. Here we briefly indicate some examples.

8.5.1. Neural Plasticity and sensory substitution. A currently very active domain of investigation in neurophysiology concerns findings which show that cortical representations of visual or somatosensory information can change as a function of stimulation, use, or lesion. For example Pascual-Leone et al. (1993) show that the sensorimotor representation of the reading finger in the cortex of proficient Braille readers becomes greatly developed at the expense of the representations of other fingers (see also Sterr et al. 1998). Sadato et al. (1998) have suggested that in proficient Braille readers, tactile processing is "rerouted" to occipital visual cortex (see also Cohen et al. 1999). The cortical representation of owl monkeys' fingertips become enlarged when the monkeys engage in haptic exploration training (Jenkins et al. 1990a). Iriki et al. (1996) found that receptive fields of bimodal (somatosensory and visual) neurons in the caudal postcentral gyrus of macaque monkeys were altered during tool use "to include the entire length of the rake or to cover the expanded accessible space." Other examples include reorganization of cortical representations as a result of intracortical microstimulation, cortical lesions, digit amputation or fusion (cf.; Jenkins et al. 1990b; Merzenich et al. 1984; 1987; Wall et al. 1986), as well as the result of von Melchner et al. (2000) showing that auditory cortex of ferrets can be "rewired" to process visual information.

8.5.2. Attention and action. Rizzolatti and his colleagues have developed a "premotor theory of spatial attention" according to which, first, "conscious space perception results from the activity of several cortical and subcortical areas, each with its own neural space representation" (Rizzolatti et al. 1994, p. 232), and second, these "neural maps" directly function in the guidance of movement and action. There are not two systems, one for spatial attention and one for action. "The system that controls action is the same that controls what we call spatial attention" (p. 256). These claims dovetail with psychophysical, psychological, and neuroscientific evidence demonstrating linkages between perception and motor action. For example, Kustov and Robinson (1996) studied "superior colliculus in monkeys as they shifted their attention during different tasks, and found that each attentional shift is associated with eye-movement preparation" (p.74). Another line of evidence linking spatial attention and motor activity comes from studies of neglect in animals and humans with damage to cortical motor areas (Kinsbourne 1987; 1995; Rizzolatti et al. 1983). Neglect appears to be best understood as a difficulty in shifting attention to the affected part of the visual field. The fact that neglect should arise from damage to cortical areas serving motor activity further demonstrates the link between attention and motor activity.

8.5.3. Two visual systems: The what and the how. In the last few years, a very influential view of the structure of the visual brain has surfaced, according to which there are two streams of visual processing, a dorsal stream and a ventral stream. Opinions differ on the exact functions of the two systems, but Ungerleider and Mishkin (1992) distinguish between a dorsal "where" system devoted to localizing objects, and a ventral "what" system devoted to identifying them. A somewhat different classification has been proposed by Goodale and Milner (1992; cf. also Milner & Goodale 1995), who emphasize that the dorsal system is concerned with coordinating actions directed towards objects, whereas in the ventral system recognition and classification operations are performed which allow persons to memorize and reason about objects. Jeannerod (1997) refers to the dorsal stream as "pragmatic," in that it provides the ability to make the necessary transformations between visual input and motor output to locate an object with respect to the body, and to grasp and manipulate it, and calls the ventral stream the "semantic" system. Evidence for this latter interpretation of the two streams hypothesis comes from studies of the effects of lesions in humans (Milner & Goodale 1995). As Milner and Goodale point out, damage to the dorsal stream is associated with impairments of visuo-motor control such as optic ataxia (Harvey 1995) in the absence of impairments of the subject's ability to make verbal reports about the shape, features, and location of what is seen. Conversely, damage to the ventral stream produces visual agnosias (Benson & Greenberg 1969; Milner et al. 1991) without impairing visuo-motor functioning.

From the standpoint of the sensorimotor contingency view we propose here, the possibility of this kind of double dissociation is not surprising. In our view, seeing is an activity depending on a broad range of capacities, for example, capacities for bodily movement and guidance, on the one hand, and capacities for speech and rational thought, on the other. To the extent that these capacities are independent, it is not surprising that they can come apart in the manner described. It is not surprising, therefore, that the dorsal system can operate in relative isolation from the ventral system.

These points lead us to doubt, on certain interpretations at least, Milner and Goodale's claim that what the visual agnosia patient DF (who retains normal visuo-motor skill) lacks is *visual awareness* of what she sees. Milner and Goodale suggest that, like DF, normals carry out visually guided actions using information that is not present in awareness; and they say that only information in the ventral stream enters awareness. According to the view developed here (the sensorimotor contingency view), people are aware of what they see to the extent that they have control over that information for the purposes of guiding action and thought. Awareness is always, we have argued, a matter of degree. Even the distracted driver is somewhat aware of what he sees, to the extent that, if we were to ask him, he would tell us what he is looking at. The case of DF is thus a case of what would seem to be partial awareness. She is unable to describe what she sees, but she is otherwise able to use it for the purpose of guiding action.

This may seem like a purely verbal dispute, but there is an important point at stake here. What makes the information conscious or aware, in our view, cannot consist just in the activity or lack of activity in a certain brain region (e.g., the ventral stream). Consciousness or awareness is not a property that informational states of the brain can just come to have in that way. Rather, visual awareness is a fact at the level of the integrated behavior of the whole organism. The work of Milner and Goodale suggests that damage to the ventral stream disrupts non-visuo-motor aspects of seeing. This is an important finding. But it would be a mistake to infer from this that the ventral stream is therefore the place where visual awareness happens.

Apart from the above provisos, the "two visual systems" view fits well with the position we develop in this paper. First, as expected from the sensorimotor contingency based approach, at the neural level there is a tight connection between seeing and moving. Second, the two-systems approach provides evidence supporting a claim we have made at different stages in this paper, namely, that seeing does not depend on the existence of unified representations of what is seen. In the two-systems approach, for example, there is not one single representation of space in the brain.

In this connection, it is worth mentioning here that the "two visual systems" approach is also relevant to the scheme for classifying sensorimotor contingencies introduced in section 2.

Processing in the two streams will in general have to be done in different coordinate systems. In order to allow reaching for an object and manipulating it, the "where" system will have to make use of the position of the object relative to the observer's body. On the other hand, in the "what" system, recognition and classification of an object will require knowledge about the intrinsic shape of the object within an object-centered coordinate system, irrespective of the observer's position with respect to it.

This distinction between an observer-centered and an object-centered coordinate system is relevant to our classification of sensorimotor contingencies. In section 2 we emphasized that there are a subset of contingencies which are due to the particular spherical structure of the eyes, to the way they move, and due to the fact that sampling is being done by means of a two-dimensional perspective projection taken at a certain distance from the object, through optics and via a retinal mosaic that have very particular properties. These visual-apparatus-dependent rules are also constrained by the fact that the objects and the eyes are embedded in three-dimensional space (rather than, say, two-dimensional space): laws of variation like the inverse square law for the amount of light reaching the eye, the linear relation between distance and projected size, will be common to all objects which are sampled through the visual apparatus.

If we were to construct a neural network connected to an eye-like input device and to a muscle-like eye-mover, and have it learn the rules of dependency between its input and output, then the first rules which the neural network would adduce would be apparatus-based rules like the ones we have just described: these rules apply in all reasonably rich environments, irrespective of the objects contained in them. Furthermore, we see that the coordinate system useful to the neural network would be the observer-based coordinate system, since what has to be learnt by the network is those (in)variance laws of the space of sensorimotor contingencies which are occasioned by the observer's own movements. This learning would also provide the system with the notion of "self," since it will allow it to distinguish between parts of its environment which it can systematically control, and parts which it cannot. The notion of "object" would also be something that would emerge from such learning: an object is something which can be removed and put back into the visual scene. These facts about what might be called object-combinatorics are independent of the identity of the objects themselves, and are related simply to their intrinsic "objectness" and their embeddedness in three dimensional space. It is possible that the dorsal "where" system could have evolved to serve this function.

The second subset of sensorimotor contingencies we referred to in our classification was the subset which we described as "object-related." These contingencies are those that allow objects to be distinguished from one another, and to be recognized independently of their position and orientation. Clearly, a neural network which was trying to adduce laws of (in)variance from sensorimotor contingencies of this kind would have an advantage in coding information in object-centered, rather than observer centered coordinates. This is known to be the case for the ventral "what" system, which could have evolved for this purpose.

Note that our classification into apparatus-related and object-related sensorimotor contingencies is a somewhat artificial division. Many of the laws underlying sensorimotor contingencies could be said to be related both to the visual apparatus and to the nature of objects. For example, the fact that objects are embedded in three dimensional space has the consequence that they can show only one face to the eye, and that as they are turned or as the observer turns around them, different parts appear and disappear. These facts are both a consequence of the fact that the eye is operating from a distance and so capturing only a single point of view – an aspect of the apparatus, – and a consequence of the fact that objects have different sides – an aspect of the objects.

8.5.4. Downward causation. There is considerable evidence that when neural correlates of consciousness have been found, they are sensitive to mood, attentional set, and task. Varela and Thompson (e.g., Thompson & Varela 2001) have referred to the modulation of individual neurons by patterns of activity of populations of neurons and also by the attitude or set of the whole animal as "downward causation." So, for example, as stressed by Varela (1984) and Varela et al. (1991, p. 93; see also Pessoa et al. 1998, p. 736; Thompson 1995, p. 217; and Thompson & Varela 2001), responses in visual cells depend on behavioral factors, such as body tilt (Horn & Hill 1969), posture (Abeles & Prut 1996), and auditory stimulation (Fishman & Michael 1973; Morell 1972). Other studies show that attention and the relevance of a stimulus for the performance of a behavioral task can considerably modulate the responses of visual neurons (Chelazze et al. 1983; Haenny et al. 1988; Moran & Desimone 1985; Treue & Maunsell 1996). Leopold and Logothetis (1999) themselves write of binocular rivalry:

> We propose that the perceptual changes are the accidental manifestation of a general mechanism that mediates a number of apparently different behaviors, including exploratory eye movements and shifts of attention. We also propose that while the different perceptions of ambiguous stimuli ultimately depend on activity in the 'sensory' visual areas, this activity is continually steered and modified by central brain structures involved in planning and generating behavioral actions. (Leopold & Logothetis 1999, p. 254)

Leopold and Logothetis suggest that to understand perceptual reversals of the kind encountered when we view an ambiguous figure, or when we undergo binocular rivalry, it is necessary to consider not only neural activity in the visual cortex, but the animal's capacities for thought and action.

8.5.5. Upshot. Work in these and other areas provides evidence in favor of ways of understanding the role of the brain in vision and consciousness that are different from the ideas in the neural correlate of consciousness and binding problem research programs. Like work in the fields of dynamic systems theory (e.g., Kelso & Kay 1987) and embodied cognition both in robots and in animals or humans (Aloimonos 1992; Bajcsy 1988; Ballard 1991; Brooks 1991; Clancey 1997; Cotterill 1995; 1997), this research suggests the importance of accounts of the brain as an element in a system, and not, as it were, as the seat of vision and consciousness all by itself.

9. Conclusion

In this paper we have put forward a new framework for thinking about the nature of vision and visual consciousness. The solution to the puzzle of understanding how consciousness arises in the brain is to realize that consciousness does not in fact arise in the brain! Visual consciousness is not a special kind of brain state, or a special quality of informational states of the brain. It is something we do.[50]

From this point of view, understanding vision amounts to understanding the various facets of the things people do when they see. We suggest that the basic thing people do when they see is that they exercise mastery of the sensorimotor contingencies governing visual exploration. Thus, visual sensation and visual perception are different aspects of a person's skillful exploratory activity (that is, exploratory activity guided by practical knowledge of the effect movement will have on nervous influx). Visual awareness depends, further, on the person's integration of these patterns of skillful exercise into ongoing planning, reasoning, decision-making, and linguistic activities. As we have argued, these ingredients are sufficient to explain the otherwise elusive character of visual consciousness.

In addition, our proposal has the advantage of providing an account of what differentiates the sensory modalities. The problem is solved naturally, without appealing to the existence of sensory-modality-specific essences or mechanisms. Just as horse riding is different from motorcycling, so is seeing different from hearing. These differences can be explained without appeal to the essences of horseback riding and motorcycle riding, and without appeal to the specific nerve energies or pathways devoted to seeing and hearing. The difference between seeing and hearing is to be explained in terms of the different things that we do when we see and hear.

These are matters of philosophical significance, and we have sought, in developing our position, to make clear that in denying the need for *qualia,* we are not denying the existence of perceptual experience, or the possibility of phenomenological reflection on experience. Rather, we have proposed a new way of thinking about what goes on when we experience – which, as we have argued throughout, captures what we believe, as experiencers, about our experiential lives, but does so in a manner that does not give rise to the mystery of the explanatory gap.

These are matters of empirical significance. The sensorimotor approach to vision we have laid out here has provided the impetus for a series of surprising experiments on what has come to be known as change blindness. The robustness of these results in turn serves to vindicate the framework itself. In addition, we have tried to demonstrate that the sensorimotor view presented here allows for a unified approach to a variety of otherwise unconnected perceptual phenomena related to, for example, perceptual completion, inattentional blindness, perceptual stability despite observer eye-movement, prosthetic perception, color vision, inverting lens adaptation, the surgical restoration of sight in the congenitally blind, blindsight, the double dissociation of optic ataxia, and visual agnosia. We think it is striking evidence of the power of the position we develop here that it is able to account for such a broad range of perceptual phenomena.

Finally, if we are correct in our analysis, there is need for reassessment, on the part of neuroscientists, of the notion of neural correlate of consciousness. A way of thinking about the neural bases of perception and action is needed that does not rest on the false assumption that the brain is the seat of consciousness. We also believe that philosophers should consider the way in which empirical results in this area suggest the formulation of a new metaphysics of mind and body.

ACKNOWLEDGMENTS

The authors deeply thank the following people for their help with different aspects of this work: Paul Bach-y-Rita, Hasna Ben Salah, Ned Block, Benédicte de Boysson Bardies, Justin Broackes, Philippe Capet, Jim Clark, Frédérique Devignemont, Don Dulany, Pedro Fonseca, Juan Gonzalez, Andrei Gorea, Stevan Harnad, Cathy Harris, Keith Humphrey, Timo Järvilehto, Pierre Jacob, Marcel Kinsbourne, Christof Koch, Stephen Kosslyn, Donald Laming, Ariane Lévy-Schoen, Michael O'Regan, Elias O'Regan, Malcolm MacIver, Erik Myin, Ernst Niebur, Luiz Pessoa, Bill Phillips, Ronan Reilly, Juan Segui, Paulo Sousa, Nigel Thomas, Jean-Pierre Rospars, Evan Thompson, Srimant Tripathy, Marc Wittman, Andreas Wohlschläger, Jeremy Wolfe, and Tiziana Zalla. Alva Noë gratefully acknowledges the support of faculty research funds granted by the University of California, Santa Cruz.

NOTES

1. Cf. also Abeles & Prut (1996); Milner (1974); von der Malsburg (1983); Kahn et al. (1997); Rodriguez et al. (1999).

2. To suppose otherwise is to assume a particular account of psychoneural correspondence. But surely, how neural processes underwrite perceptual experiences is precisely what requires explanation. We can frame the problem in terms of Teller's notion of linking propositions (Teller 1983; 1984; Teller & Pugh 1983). Linking propositions specify the mappings between the forms of neural responses and the qualities of percepts.

3. See Gorea (1991) for a discussion of the relation of modern psychophysics to Müller's concept.

4. Koenderink (1984b) gave an example of how the shape of a tomato changes as you look at it from different angles. He called this the "aspect graph" or "visual potential" of the tomato. The concept of aspect graph has since then been extensively investigated in artificial vision.

5. This classical reasoning is an instance of what Pessoa et al. (1998) call "analytic isomorphism," that is, the view that at the neural substrate of an experience there must be an isomorophism between percept and substrate. Analytic isomorphism comes up again in our discussion of the neural basis of vision in section 8.

6. Koenderink (1984a) has a very perspicacious discussion of what it is to perceive, rather than simply to record information, where he makes this point.

7. Heil (1983), agreeing with Gibson and rejecting Müller's idea of physiological "channels" associated with different senses, also attempts a taxonomy of the different senses but does not suggest the idea that it could be the laws obeyed by the sensorimotor contingencies that are the essential fact that differentiates them.

8. Note that it could be claimed that Müller's idea of specific nerve quality could be salvaged by supposing that what differentiates the senses is different *calculations* that are done in the different pathways. This was suggested by Wittmann, Pöppel, and Schill, reviewers of the original version of this manuscript. In a way this is what is being proposed by the present approach, although we emphasize that the calculation itself is not enough. What is needed is for the structure of the input/output relationships to obey different laws.

9. Note that we have been careful not to say that vision or horseriding *provide* different experiences: the experience *is the fact of* engaging in the activities. The activities, we claim, are not *providing* an experience – though people often use the word provide in this way, we claim this is a figure of speech, and not indication of a true experience-generating mechanism. It is precisely this kind of misunderstanding which gives rise to the problem of the explanatory gap. Cf. section 6.3.

10. The possibility of machine awareness raises issues that go beyond the scope of the present discussion. We note here that because we admit that awareness comes in degrees, we are willing to say that to the extent that machines can plan and have rational behavior, precisely to that same extent are they also aware. But clearly, given the limitations of current machines' planning and rational behavior, and given the lesser diversity of their environmental interactions, the degree of awareness will be accordingly limited. If a chess-playing machine were able to purposefully lose a game so as to avoid upsetting a child, or if a medical diagnosis system were able to lie to a cancer patient about his condition, we would be more willing to accord higher degrees of awareness to it.

11. Merleau-Ponty (1968) has also compared vision to palpation.

12. We unfortunately became aware of Järvilehto's work too late to be able to give it full consideration.

13. Järvilehto (1999) has also made this point.

14. It is often assumed that transients must necessarily direct attention to a *location.* But presumably location is only one feature of visual stimuli, and in the brain, location may have a similar status to other features, like color, orientation, contrast, etc. Could it be that attention can be directed to aspects of a stimulus defined by such other features? For example, is it possible to direct attention to all the red items in a scene, or to scene region constituted by a 3D surface? Cf. Pylyshyn (1988) on this issue.

15. An excellent discussion of these topics can be found in Thomas (1999), who makes a convincing argument in favor of an "active perception" approach very similar to ours.

16. As observed by Stephen LaBerge (personal communication).

17. They observed no eye movement advantage but less than 50% correct performance in counting a grating pattern of identical vertical bars: it may be that the observers were using a strategy of estimating the number of bars by evaluating the number on the basis of the overall width of the pattern.

18. A portion of the existing data purportedly measuring the extraretinal signal under conditions of normal viewing can, to some degree, probably be explained by assuming that they are due to purely retinal effects (smear, retinal persistence, differences in spatio-temporal effects in central and peripheral vision; cf. O'Regan 1984).

19. It could be argued that people actually do have a detailed, picture-like internal representation of the outside world, but that it is destroyed at each saccade or on interruption by flicker and other transients. Alternately, as suggested by reviewers Wittmann, Schill, and Pöppel of our manuscript, it may be that we deceive ourselves as to the amount of detail we think we see in the representation. Such arguments are hard to square with data showing interaction of the change blindness effects with central/marginal interest manipulations, and with the data from the "mudsplash" experiments, among others. (cf. O'Regan et al. 1996; Rensink et al. 1997; 2000). Similar alternatives have also been discussed by Simons (2000b).

20. This view of the phenomenology of color perception is related to the idea of D'Zmura and Lennie (1986), who suggest that nonhomogeneity in retinal cone distributions could indeed be made use of by the visual system to determine surface reflectances. But in general, most current views of color perception assume that perceived color derives from applying some kind of color constancy calculation to the output of the long, medium, and short wavelength cone channels. The idea that perceived color is not the output of a constancy calculation, but rather is *constituted* by the applicability of laws of variation under eye movements, lighting conditions, and surface movements, appears not to have been seriously investigated up to now.

21. Broackes additionally notes: "And if it is puzzling how a dynamic property can make itself manifest in a static perception ('how can a disposition to present a variety of appearances be visible in a single appearance?'), then we already have, in familiar discussions of aspect-shift, the theoretical apparatus for a solution. It is because there is 'the echo of a thought in sight'." Broackes quotes Strawson (1974, pp. 52–53) who says: "To see [a newly presented object] as a dog, silent and stationary, is to see it as a possible mover and barker, even though you give yourself no actual images of it as moving and barking."

22. Broackes says that contrary to what he said in Broackes (1992), he is either protanomolous or protanope. (Broackes, personal communication.)

23. Cole (1991) has also invoked these studies in a functionalist defense against the inverted spectrum problem.

24. Curiously, many people wearing normal glasses seem to voluntarily peer over the rims of their glasses when they look at you, as though this procured some kind of advantage in seeing.

25. Dolezal stresses that the use of the terms upside down and right side up is confusing, and guards against saying that the world comes again to appear right side up. He says that in his experiment the final state of adaptation could be distinguished from the state before the experiment. Part of the reason for this could be that the duration of the adaptation was perforce limited, and because use of inverting goggles necessarily involves other constraints like the limited field of view and the weight of the apparatus. Howard and Templeton (1966) also stress the need to be wary of the terms upside down and right side up (see also Linden et al. 1999; Smith & Smith 1962).

26. Chapter 8 in Taylor's book contains a detailed, behaviorist theory of the effects of inversion of the visual world, referring to specific results of Stratton, Ewert, and Kohler. The outcome appears to be that the observed adaptation effects are to be expected, and that the nativist theory is "shattered" (p. 168).The chapter includes a mathematical appendix by Seymour Papert, who was the subject in Taylor's left-right inversion experiment.

27. The situation may be similar to what happens when you move to a new town, and attempt to orient yourself. It takes some time before local and global landmarks merge into a coherent representation of the town. Until that happens (and it may never do), you may make gross mistakes. For example, you may be perfectly able to orient yourself locally, but be unable to correctly indicate the direction of a well-known global landmark.

28. Bedford (1995) has a theory of perceptual learning which is related to the theory presented here.

29. Movies of the demonstrations can be found on the first author's web page at http://nivea.psycho.univ-paris5.fr. See also http://www.wjh.harvard.edu/~viscog/change/demolinks.shtml.

30. Movie demonstrations of some film-cut and other effects can be found on Simons' page at http://coglab.wjh.harvard.edu.

31. The following quote from Haines (1991) is an example: "Pilot F was a high-flight-time Captain who demonstrated exceptionally good performance both with and without HUD. The runway obstruction run was his seventh data run. He indicated his 'Decision (140 ft) . . to land (110 ft)', and proceeded to do so. The experimenter terminated the run at an altitude of 50 ft. The pilot was surprised. Captain: 'Didn't get to flare on this one.' First Officer: 'No you didn't . . . I was just looking up as it (the picture) disappeared, and I thought I saw something on the runway. Did you see anything?' Captain: 'No, I did not.' The experimenters suggested that an equipment failure was probably to blame. Both of these pilots saw the obstruction during the second exposure without HUD (13 runs and 21 runs later, respectively) and executed missed approaches. Later, when he was shown the videotape of this run, Pilot D said, 'If I didn't see it (the tape), I wouldn't believe it. I honestly didn't see anything on that runway'."

32. In Noë & O'Regan (2000) we discuss some philosophical aspects of the inattentional blindness work.

33. This demonstration may not work if the file is being viewed on the web or has been printed with the option of substitution of typography enabled. The point is that there are two ways of forming an "a"; one similar to the hand-written ɑ (a circle with a line next to it), and one similar to a typewriter a. If hand-written-like a's are mixed into a text, provided they have the same height and density as normal a's, this will generally not be noticed.

34. The word "of" is repeated. Repetitions of the word "the" can also be easily missed.

35. There are nine f's. Many people fail to count the f's in the three occurrences of the word "of."

36. In fact sensation itself is an abstraction, as already noted by James (1890/1950, vol. 2, p. 3).

37. Humphrey and Humphrey (1985) quote a blind man (D. Lepofsky, 1980) who has used a binaural sonic sensor mounted on eyeglasses for 5–10 hours a week for three years: "I am at the point that I react very naturally to its signals. I no longer have to think about what each signal could mean, rather, I react instinctively. I go around someone on the sidewalk without even realizing I've done it: that's how much a part of you it becomes."

38. Lenay et al. (1999) have discussed other reasons why TVSS systems have met with less enthusiasm on the part of blind people than might have been expected. He says: "Ce que cherche l'aveugle qui accepte de se plier à l'apprentissage du dispositif de couplage, c'est d'avantage la connaissance de ce dont les voyants lui parlent tant : les merveilles du monde visible. Ce qu'il espère, c'est la jouissance de cette dimension d'existence qui lui est inconnue. Or, ce n'est pas ce que donne ces dispositifs. Il y a de fait, de nombreuses différences entre le couplage artificiel et notre couplage visuel: il n'y a pas de couleur, peu de points, une caméra dont les mouvements sont difficiles et limités, ce qui donne une grande lenteur à la reconnaissance de la situation. Ce couplage sensori-moteur ressemble bien par certains aspects à celui de notre vision, mais l'expérience qu'il permet est toute différente, comme peuvent d'ailleurs bien le comprendre les voyants qui se prêtent à son apprentissage. Le dispositif de Bach-y-Rita ne réalise pas une substitution sensorielle, mais une addition, l'ouverture d'un nouvel espace de couplage de l'homme avec le monde."

39. Howells (1944) cited by Taylor (1962, p. 246) is an interesting example where association of a low and high pitched tone with red and green respectively, over 5,000 trials, gave rise to a perception of white being tinged with red and green when white was associated with the tones.

40. Bedford (1995) has a theory bringing together the McCollough effect and adaptation to prism displacement which is similar in concept to the present theory.

41. Of course, this is not to deny that vision may, under certain circumstances, involve feelings or sensations of a non-visual nature. So, for example, if you are trying to track the movement of an object without moving your head, you may feel a certain distinctive eye strain. If you witness an explosion, you may feel dazzled in a way which causes definite sensations in the eyes. If vision is, as we have argued, a mode of activity, then there may be all sorts of features that the activity consists of which in this way contribute to its "felt character." But crucially these are not *intrinsic* or *defining* properties of the experiencing, that is, they are not what philosophers think of as qualia. They are rather more or less accidental accompaniments of the activity of seeing on a particular occasion. Note, similar points can be made for the other sensory modalities. Bach-y-Rita (1996) has noted that perceptual experience may have a qualitative aspect in yet another sense. For those capable of vision, certain experiences may have a definite *affective* quality. This affective quality was reported to be absent in the quasi-visual experiences of patients using TVSS. So, for example, such patients lacked the familiar "feel" of emotion and familiarity when looking at a picture of a loved one, or the erotic charge that may be delivered by certain images in normal perceivers. Bach-y-Rita reports these differences as differences in qualia. This usage differs from that in the philosophical literature. In any event, we do not deny that experiences may be associated in this way with affect. In fact, the sensorimotor contingency view offers a basis from which to explain what may be going on here. One might speculate that what prevents tactile visual experiences from acquiring a full affective charge is the fact that tactile vision is not perfectly mastered, that is to say, it is not fully integrated into a sensorimotor skill set. A direct consequence of this strangeness is the fact that one's intimate dealings with one's loved ones have not been mediated by the exercise of the relevant sensorimotor skills.

42. Of course, when you drive a Porsche for the first time, you may at first lack *confident knowledge* of how the car will respond to your actions. Insofar as you are an experienced driver of cars, you will exercise confident mastery of how to drive. In so far as you are new to Porsches, you may be tentative and exploratory. You try to learn how the car performs. The distinctive feel of driving a Porsche for the first time thus can be understood to differ from the experience of the connoisseur.

43. One of the hallmarks of the tradition known as Phenomenology, associated with the work of Husserl and Merleau-Ponty, is a clear and rigorous conception of the methodology of first-person investigations of experience. Of great importance for appreciating this tradition, Husserl and Merleau-Ponty make contributions toward the development of a first-person study of consciousness which does *not* rely on the problematic conception of qualia criticized above. We are broadly sympathetic to work in this tradition. For recent contributions, see Varela and Shear (1999) and Petitot et al. (1999). Other traditions may also provide methods and concepts for first-person investigations of experience, for example, the mindfulness-awareness tradition in Tibetan Buddism. See Varela et al. (1991). Our use of the term "phenomenology" above, however, is not meant to refer specifically to these traditions but, rather, to the general problem to the solution of which these traditions make a contribution. Our central aim above is to make clear that we do not believe that there is any incompatibility between the sensorimotor contingency theory and a more full-blooded phenomenological project.

44. This has been shown for example by Chun and Nakayama (2000) in the context of experiments on change blindness (cf. also Chun & Jiang 1998).

45. An interesting question arises about the relation between what we think of as the animal's active engagement with the world and what, in the Phenomological Tradition, is known as the lived-body. This is a subject for further inquiry.

46. Research in the Phenomenological Movement associated with Husserl and Merleau-Ponty is concerned precisely with the development of just such first-person methods. It is ironic that Dennett criticizes Phenomenology from the heterophenomenological perspective (see, e.g., Dennett 1991, p. 44), since, as noticed by Thompson et al. (1999) and also Marbach (1994), Dennett's misdescriptions of experience often turn on misunderstandings that have been clearly understood as such within the Phenomenological Tradition.

47. In a criticism of Crick and Koch's arguments, Cogan (1995) also suggests that the notion of "perceptual moment" may not be useful. Dennett (1991) notes a similar point in his "multiple drafts" theory of consciousness.

48. A similar argument was made in section 5.5 with regard to the "filling in" of the blind spot: there may actually be what look like filling in processes in the brain, but these cannot be what provide us with the impression of the blind spot being filled in.

49. Of course, it is possible by attending to blinks (or eye movements), to become aware of the change in sensory input that they cause. But normally people do not attend to blinks or eye movements, and do not notice them. Certainly they do not attribute the sensory interruptions they cause, to changes in external objects.

50. Varela et al. (1991) and Thompson et al. (1992) also make this point. In their terminology, consciousness is something we enact.

Open Peer Commentary

Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.

Editorial commentary

Let us simplify the problem of "consciousness" or "visual consciousness": Seeing is feeling. The difference between an optical transducer/effector that merely interacts with optical input, and a conscious system that sees, is that there is something it *feels like* for that conscious system to see, and that system feels that feeling. All talk about "internal representations" and internal or external difference registration or detection, and so on, is beside the point. The point is that what is seen is felt, not merely registered, pro-