Is Seeing All It Seems? Action, Reason and the Grand Illusion

Andy Clark

Cognitive and Computing Sciences

University of Sussex

Brighton, UK

andycl@cogs.susx.ac.uk

Final Draft:

Abstract

We seem, or so it seems to some theorists, to experience a rich stream of highly detailed information concerning an extensive part of our current visual surroundings. But this appearance, it has been suggested, is in some way illusory. Our brains do not command richly detailed internal models of the current scene. Our seeings, it seems, are not all that they seem. This, then, is the Grand Illusion. We think we see much more than we actually do. In this paper I shall (briefly) rehearse the empirical evidence for this rather startling claim, and then critically examine a variety of responses. One especially interesting response is a development of the so-called 'skill theory', according to which there is no illusion after all. Instead, so the theory goes, we establish the required visual contact with our world by an ongoing process of active exploration, in which the world acts as a kind of reliable, interrogable, external memory (Noe, Pessoa and Thompson (2000), Noe (2001). The most fully worked-out versions of this response (Noe and O'Regan (2000), O'Regan and Noe 2001) tend, however, to tie the contents of conscious visual experience rather too tightly to quite low-level features of this ongoing sensorimotor engagement. This (I shall argue) undervalues the crucial links between perceptual experience, reason and intentional action, and opens the door to a problem that I will call 'sensorimotor chauvinism': the premature welding of experiential contents to very specific details of our embodiment and sensory apparatus. Drawing on the dual visual systems hypothesis of Milner and Goodale (1995), I sketch an alternative version of the skill theory, in which the relation between conscious visual experience and the low-level details of sensorimotor engagement is indirect and non-constitutive. The hope is thus to embrace the genuine insights of the skill theory response, while depicting conscious visual experience as most tightly geared to knowing and reasoning about our world.

I.   Amazing Card Tricks.

There is an entertaining web site[i] where you can try out the following trick. You are shown, on screen, a display of six playing cards (new ones are generated each time the trick is run). In the time-honoured tradition, you are then asked to mentally select and recall one of those cards. You click on an icon and the cards disappear, to be replaced by a brief 'distracter' display. Click again and a five card (one less) array appears. As if by magic, the very card that you picked is the one that has been removed. How can it be? Could the computer have somehow monitored your eye movements?

I confess that on first showing (and second, and third) I was quite unable to see how the trick was turned. It works equally well, to my surprise, using OHP's or a printout!

Here's the secret. The original array will always comprise six cards of a similar broad type e.g. six face cards, or six assorted low-ranking cards (between about 2 and 6) etc. When the new, 5 card array appears, NONE of these cards will be in the set. But the new 5 card array will be of the same type e.g. face cards, low cards, whatever. In this way, the trick capitalises on the visual brain's laziness (or efficiency, if you prefer). It seems to the subject exactly as if all that has happened is that one card (the one they mentally selected!) has gone from an otherwise unchanged array. But the impression that the original array is still present is a mistake, rooted no doubt in the fact that all we had actually encoded was something like 'lots of royal cards including my mentally selected king of hearts'[ii]

Most magic tricks rely on our tendency to overestimate what we actually see in a single glance, and on the manipulation of attention so as to actively inhibit the extraction of crucial information at certain critical moments. Las Vegas and the Grand Illusion go hand in hand.

Daniel Dennett makes a similar point using a different card trick. He invites someone to stand in front of him, and to fixate his (Dennett's) nose. In each outstretched arm Dennett holds a playing card. He brings his arms in steadily. The question is, at what point will the subject be able to identify the colour of the card? Here too, we may be surprised. For colour sensitivity, it turns out, is available only in a small and quite central part of the visual field. Yet my conscious experience, clearly, is not as of a small central pool of colour surrounded by a vague and out of focus expanse of halftones. Things look coloured all the way out. Once again, it begins to look as if my conscious visual experience is overestimating the amount and quality of information it makes available. Talk of a Grand Illusion[iii] is clearly on the cards.

II.   Seeing, Seeming and The Space for Error

How should we characterise the kind of visual overestimation highlighted by the card tricks (and by the experimental evidence to be examined in the next section)? The matter is delicate. We cannot, surely, make much sense of the idea that we are wrong about how our visual experience visually *seems*. If it seems to me as if I see colours 'all the way out' then that simply *is* how it seems to me: there is little space for error in this space of seemings.

About what, then, might I actually be mistaken? Not about the visual seeming itself. And not, of course, about the actual real-world scene. That scene, in the typical case, really is coloured all the way out, and really is rich in detail etc. The space for genuine error is thus rather small. It must centre on what we *come to believe* as a result of how our visual experience presents the world to us. Perhaps, for example, we come to believe that our brains are constructing, moment-by-moment, a richly detailed, constantly updated internal representation of the full, and fully coloured, visual scene. Noe, Pessoa and Thompson (2000) term this the 'reconstructionist' model of vision. The idea of a rich visual buffer (Feldman 1985), in which more and more information accumulates over time is also grist to this kind of mill.

On both these counts, science could easily show us to be wrong. But (As Noe et al point out) our error would be a technical one: an error in the theory that our experience leads us (some of us) to construct. This sounds rather less grand than the claim that we are simply mistaken about the nature of our own visual experience, or subject to some kind of experiential illusion.

It is possible, even more radically, to be sceptical about the very idea of a 'way things visually seem to us', at least insofar as such seemings are depicted as objects of conscious awareness. In this vein Mark Rowlands (ms) suggests that "what it is like to undergo an experience is not something *of* which we are aware but something *with* which we are aware in the having of an experience". Visual seemings, he suggests, are not objects of normal visual experience, so much as modes of experiencing the world. As such, they do not seem any way at all: instead, via the experiences, the *world* seems this way or that.

There is good reason, then, to be a little cautious of statements such as the following:

> The visual world seems to naïve reflection to be uniformly detailed and focussed from the centre out to the boundaries, but…this is not so.
> Dennett 1991 p.53

Much depends, of course, on just what gets built into the idea of 'naïve reflection' (how naïve, and by whom?). At the very least it looks likely that the path to Grand Illusion is paved by inferences: inferences that concern the internal machinery of seeing and take us far beyond the simple act of visually knowing the world.

But surely, someone will reply, there is also an illusion within the domain of the experience itself. It does not seem to us as if our colour vision is as restricted as it is. So there is error in the way things visually seem. Rowlands would reject this, for the reasons just examined. But in any case, the response assumes, illegitimately, that

whatever is true of our experience must be true of the underlying machinery, and at a kind of instantaneous time-slice at that. Perhaps there need be no such match. Or if there is a match, it may be between visual activity over time and the contents of the experience.

The space for genuine error, I conclude, is not as large as it may initially appear. The kernel of truth in the sweeping talk of a Grand Illusion must be sought in a careful analysis of certain theoretical commitments. With that in mind, let's start by taking a look at the experimental data that fuels much of the debate.

III. What Goes Unnoticed[iv]

It is well known that the human visual system supports only a small area of high-resolution processing; an area corresponding to the fraction of the visual field which is currently foveated. When we inspect a visual scene, our brains actively move this small high-resolution window around the scene, alighting first on one location, then another. The whole of my bookcase, for example, cannot possibly fit into this high-resolution window at a glance, at least while I remain seated at my desk. My overall visual field (including the low-resolution peripheries) is, of course, much larger, and a sizeable chunk of my bookshelf falls within my course-grained view. As long ago as 1967[v] it was known that the brain makes intelligent use of the small high-resolution area, moving it around the scene (in a sequence of so-called "visual saccades") in ways suited to the specific problem at hand. Human subjects confronted with *identical* pictures, but preparing to solve *different* kinds of problem (e.g. "give the sex and ages of the people in the picture," "describe what is going on" and so on) show very different patterns of visual saccade. These saccades, it is also worth commenting, are fast – perhaps three per second – and often repetitive, in that they may visit and re-visit the very same part of the scene.

One possibility, at this point, was that each saccade is being used to slowly build-up a detailed internal representation of the *salient* aspects of the scene. The visual system would thus be selective, but would still be using input to build up an increasingly detailed neural image of (selected aspects of) the scene. Subsequent research, however, suggests that the real story is even stranger than that.

Imagine that you are the subject of this famous experiment.[vi] You are sat in front of a computer screen on which is displayed a page of text. Your eye movements are being automatically tracked and monitored. Your experience, as you report it, is of a solid, stable page of text which you can read in the usual way. The experimenter then reveals the trick. In fact, the text to the left and right of a moving 'window' has been constantly filled with junk characters, not proper English text at all. But because the small window of normal, sensible text has been marching in step with your central perceptual span, you never noticed anything odd or unusual. It is as if my bookshelf only contained (at any one moment) four or five proper, clearly titled books, and the rest was fuzzy, senseless junk. But those four or five proper items were moved about as my eyes saccaded around the scene! In the case of the screen of text, the window of "good stuff" needed to support the illusion is about 18 characters wide, with the bulk of those falling to the right of the point of fixation (because English is read left to right).

Similar experiments[vii] have been performed using pictures of a visual scene, such as a house, with a parked car and a garden. As before, the subject sits in front of a computer generated display. Her eye movements are monitored and, while she saccades around the display, changes are clandestinely made: the colours of flowers and cars are altered, the structure of the house may be changed. Such changes, likewise, go undetected. We now begin to understand why the patterns of saccade are not cumulative – why we visit and repeatedly re-visit the same locations. It is because our brains just don't bother to create even the kind of selective-but rich inner models we earlier considered. Why should they? The world itself is still

there, a complex and perfect store of all that data, nicely poised for swift retrieval as and when needed by the simple expedient of visual saccade to a selected location. The kind of knowledge that counts, it begins to seem, is not detailed knowledge of what's out there, so much as a broad idea of what's out there: one capable of informing those on-the-spot processes of information retrieval and use.

Finally, lest we suspect that these effects (known as "change blindness") are somehow caused by the unnaturalness of the experimental situation, consider some recent work by Dan Simons and Dan Levin. Simons and Levin (1997) took the research into the real world. They set up a king of slapstick scenario in which an experimenter would pretend to be lost on the Cornell Campus, and would approach an unsuspecting passer-by to ask for directions. Once the passer-by started to reply, two people carrying a large door would (rudely!) walk right between the enquirer and the passer-by. During the walk through, however, the original enquirer is replaced by a different person. Only 50% of the direction-givers noticed the change. Yet the two experimenters were of different heights, wore different clothes, had very different voices and so on. Moreover, those who did notice the change were students of roughly the same age and demographics as the two experimenters. In a follow-up study, the students failed to spot the change when the experimenters appeared as construction workers, placing them in a different social group. The conclusion that Simons and Levin (1997, p.266) draw is that our failures to detect change arise because "we lack a precise representation of our usual world from one view to the next". We encode only a kind of 'rough gist' of the current scene – just enough to support a broad sense of what's going on *insofar as it matters to us*, and to guide further intelligent information-retrieval as and when needed.

In all these cases, the unnoticed changes are made under cover of some distracting event: they are made during a saccade, a screen flicker, a movie cut and so on. These mask the visual transients (motion cues) that might otherwise draw attention to the fact that something is changing.

The importance of attention is underlined by a different series of experiments due to Mack and Rock (1998). These concern what they call 'inattentional blindness'. The focus here is not on change over time, so much as upon what can be noticed in a static scene (for an excellent discussion of this at times elusive distinction, see Rensink (2000)). The question guiding the experiments was thus simply "What is consciously perceived in the absence of visual attention?" and more particularly "Will an object, unexpectedly presented in the visual field, tend to be consciously noticed?"

A typical experiment went like this. Subjects were shown a visually presented cross (on a computer screen) and asked to report which arm of the cross was longer. The difference was small, so the task required some attention and effort. The cross was briefly presented (for about 200ms) and then a mask (an unrelated, patterned stimulus) shown. Then the subjects made their reports. On the third or fourth trial, however, a 'critical stimulus' was also shown on the screen with the cross. It might be a coloured square, a moving bar, and so on. Subjects were not expecting this. The question was, would it be consciously noticed?

The experiment was run in two main forms. In the first, the cross was presented centrally, at the point of fixation, and the critical stimulus parafoveally (to the side). In the second, subjects fixated a central point, the cross was presented parafoveally and the critical stimulus appeared just beside the fixation point. With the critical stimulus presented parafoveally, 25% of subjects failed to spot it. This is already a surprising result. But when presented near fixation, a full 75% of subjects failed to report the stimulus! Why the difference? Perhaps the need to focus visual attention away from the normal central point demanded increased visual effort and attention. Also, subjects may have had to actively inhibit information from the point of fixation. Interestingly, in both cases, subjects did spot more meaningful stimuli, such as their own names, or a smiley face: a quirk which will turn out to make good sense in the light of our final story.

Our unconscious and inattentive use of visual input may, in addition, be surprisingly extensive. For example, other words which were used as the critical stimulus in some of Mack and Rock's experiments, though unnoticed, were capable of priming subsequent choices. Exposure to the word 'provide' increases the likelihood of the stem completion 'pro' with 'vide', despite the subjects total lack of conscious awareness of the initial stimulus.

From all of this, Mack and Rock draw a strong and unambiguous conclusion. There is, they claim "no conscious perception at all in the absence of attention" (op cit. p.227). This would be trivial if attention itself were defined in terms of, say, our conscious awareness of an object. But what Mack and Rock really mean is that there is no conscious perception in the absence of expectations and intentions directed at an object. They offer no clear definition of attention itself. But *inattention* is quite well characterised:

> For a subject to qualify as inattentive to a particular visual stimulus, the subject must be looking in the general area in which it appears, but must have no expectation that it will appear nor any intention regarding it.
> Mack and Rock (1998) p.243

The importance of attention and expectation is nowhere more apparent than in another famous experiment[viii] in which subjects watch a video of two teams, one in white and one in black, passing basketballs (one ball per team). The viewer must count the number of successful passes made by the white team. Afterwards, subjects are asked whether they saw anything else, anything unusual. In fact, about 45 seconds into the film an intruder walks through the players. The intruder might be a semi-transparent, ghostly figure of a woman holding an umbrella, or a semi-transparent gorilla (without any umbrella). Or even, on some trials, a fully opaque woman or gorilla! In the semi-transparent condition, 73 % of subjects failed to see the gorilla, and even in the opaque condition, 35% of subjects failed to spot it (see Simons (2000) p.152).

Simons interprets these results as suggesting the possibility:

> That our intuitions about attentional capture reflect a *metacognitive error*: we do not realise the degree to which we are blind to unattended and unexpected stimuli and we mistakenly believe that important events will automatically draw our attention away from our current task or goals.
>
> Simons (2000) p.154 My emphasis.

It is easy to see, given the work on change blindness and inattentional blindness, why talk of a Grand Illusion can seem so compelling. Conscious vision, it can quickly seem, delivers far less than we think.

IV. Diagnoses

There are, as far as I know, four main responses to the bodies of data reviewed in section 2. They are:

i. The Grand Illusion
ii. Fleeting Awareness with Rapid Forgetting
iii.Projected (memory-based) Richness
iv.Skill Theory

Hints of the Grand Illusion response can be seen in many treatments from Dennett 1991 onwards, including Ballard (1991), O'Regan (1992) Churchland et al (1994), Clark (1997) and Simons and Levin (1997). The idea is simple and attractive. We do indeed (it is claimed) seem to experience a continuous stream of richly detailed, wide-angled, fully coloured, new-input-sensitive information in the conscious visual modality. But the seeming is just that: a seeming. It is an illusion caused by our ability to visually visit and re-visit different aspects of the scene according to our projects and as 'captured' (sometimes) by motion transients etc. We thus think that our at-a-glance visual uptake is much richer than it is due to our active capacity to get more information as and when required. Thus we read that:

> The feeling of the presence and extreme richness of the visual world is…a kind of illusion, created by the immediate availability of the information in (an) external store [the real world]
> O'Regan (1992) p.461

> The experiential nature of the visual scene is a kind of subjective visual illusion created by the use of rapid scanning and a small window of resolution and attention.
> Clark (1997) p.31

> The visual system provides the illusion of three-dimensional stability by virtue of being able to execute fast behaviours.
> Ballard (1991) p.60

The real, non-illusory, knowledge built up by our ongoing visual contact with a scene is, on these models, quite schematic and high-level. We maintain a general sense of the situation, just enough to guide attention and saccades while we actively engage in a scene-related task.

An alternative hypothesis is the so-called "fleeting awareness" account (also known as "inattentional amnesia") presented by Wolfe (1999). The suggestion is that our moment-by-moment conscious visual experience may be rich and detailed indeed, but that we simply forget, pretty well immediately, what the details were, unless they impact our plans and projects very directly. Since these paradigms always involve questioning at least fractionally after the event, subjects say they did not see the new objects etc. But this reflects a failure of memory rather than a deficit in ongoing conscious visual experience.

Some element of forgetting may, I accept, be involved in some of these cases. But overall, the hypothesis strikes me as unconvincing. First of all, it is not really clear whether 'seeing-with-immediate-forgetting' is really any different from not seeing at all. (Recall

Dennett's 1991 discussion of Stalinesque versus Orwellian accounts). Second, we know that only a very small window of the visual field can afford high resolution input (Ballard 1991), and we know that attentional mechanisms probably limit our capacity to about 44 bits (plus or minus 15) per-glimpse (see Verghese and Pelli (1992), Churchland et al (1994). So how does all that fleeting richness get transduced? And lastly, as Simons (2000) nicely points out, the inattentional amnesia account seems especially improbable when the stimulus was an opaque gorilla presented for up to 9 seconds. Could we really have been consciously aware of that and then had it slip our minds?

A third diagnosis invokes memory in a more active role: in the role of 'filling in' the missing detail. The suggestion is that our conscious visual experience is enriched 'top down' by stored memories and expectations. So we do indeed see a highly detailed scene. It is just that, in a sense, we make most of it up! I suspect that this suggestion contains an important kernel of truth, and we shall return to it in section VI, where we display strong links between conscious experience and certain kinds of memory system.

By far the most interesting, deep and challenging response, however, is one which rejects the Grand Illusion diagnosis while nonetheless *accepting* the poverty of the moment-by-moment internal representations that the visual system creates and maintains. According to this response, the Grand Illusion is itself a chimera, caused by the fall-out from a classical, disembodied approach to perception. If we were to *really* embrace the idea of cognition as the active engagement of organism and world, the suggestion goes, we would see that there is no Grand Illusion after all. Hints of this idea were present, right alongside the Grand Illusion diagnosis, in O'Regan (1992), drawing on MacKay (1967). But the most clear-cut, powerful and persuasive versions are those of Noe, Pessoa and Thompson (2000), Noe (2001), Noe and O'Regan (2000), and O'Regan and Noe (In Press-2001).

Before proceeding, I must enter a caveat. Noe, Pessoa and Thompson (2000) argue, convincingly I believe, that the Grand Illusion diagnosis is a mistake, and that it is a mistake caused by failing to appreciate that seeing is a temporally extended process involving active exploration of the environment. My critical concern, in what follows, is not with this general claim but with the specific way it is unpacked, in the context of a more fully worked-out version of the skill-theory, in O'Regan and Noe (2001). For this specific version of the skill-theory (I shall argue) ties conscious visual experience too closely to the full gamut of (what one might question-beggingly describe as) the 'implementation detail' of the visual apparatus. My goeal will be to develop an account in the spirit both of skill-theory and of Noe, Pessoa and Thompson's critique of the Grand Illusion claim. But it will be an account that leaves room for some details of the visual apparatus to make no difference to the contents or character of conscious visual experience.

A good place to start is with the MacKay-based example given by O'Regan (1992) (and mentioned in O'Regan and Noe (In Press-2001)). The reader is invited to consider the tactile experience of holding a bottle in the hand. As you hold the bottle, your fingertips are in touch with just a few small parts of the surface. Yet what you experience is having the whole bottle in your grasp. This, it is argued, is because:

> My tactile perception of the bottle is provided by my exploration of it with my fingers, that is, by the sequence of changes in sensation that are provoked by this exploration and by the relation between the changes that occur and my knowledge of what bottles are like…I expect that if I move my hand up…I will encounter the cap or cork..
> O'Regan (1992) p.471, following MacKay (1967)

Our conscious tactile experience as of holding a whole bottle is thus generated by our implicit (not conscious, propositional) knowledge of how those more local finger-tip sensations would flow and alter were

we to actively explore the surface. The conscious perceptual content is thus based on actual and potential action cycles rather than on the instantaneously transduced information. This kind of implicit knowledge of reliable flows of sensory input during the execution of movements and actions is what O'Regan and Noe (In Press-2001) dub "mastery of laws of sensorimotor contingency".

Consider next the case of conscious seeing. In some ways, this is an even better case for the sensorimotor contingency model, since here we combine the input from the high resolution moveable fovea with low resolution peripheral signals capable of further aiding intelligent exploration. Our visual awareness of the scene before is thus grounded in a potent combination of:

i.   Our implicit knowledge of how the foveated input will change as we actively explore the scene
ii.   The ongoing sequence of cues provided by peripheral pick-up
iii.   Quite high-level knowledge of the nature of the scene or event we are witnessing.

Taking all these into account, it does indeed seem churlish to describe our ongoing visual experience as misleading. The impression we have of rich and available detail is correct, as long as we avoid a kind of temporal error. The Grand Illusion diagnosis trades, perhaps illegitimately, upon the idea that the content of conscious visual perceiving is given by some instantaneous, fully internally represented, deliverance of the sense organs. It trades upon the idea of simple inner state, without past or future trajectory.

It may perhaps be useful to consider an analogy. When you encounter certain web pages, you may have a strong impression of richness. This impression is grounded in your perception of a screen rich in pointers to other sites, and your implicit knowledge that you can access those other sites with a simple flick of the mouse. Such a web page leaves us poised to access a wealth of other data pretty much at will.

Following Kirsh (1991) I have argued elsewhere (Clark (1993)) that externally stored information which is poised for swift, easy and intelligent retrieval as-and-when needed should sometimes be regarded as *already represented* within the cognitive system. (It is this general commitment that leads, for example, to the 'extended mind' story found in Clark and Chalmers (1998)). The real-world scene, as O'Regan and others have pointed out, often meets this criterion, and should thus be regarded as a temporary, ever-changing module of external memory. The act of foveation-with-attention effectively moves information out of this module and into a kind of working memory buffer, making it available for the guidance of intentional and deliberate action (see Ballard et al (1997) on 'deictic pointers' for a worked out, experimentally-supported version of this kind of story). Indeed, as long ago as 1972 Newell and Simon commented that:

> From a functional viewpoint, the STM should be defined not as an internal memory but as the combination of (1) the internal STM and (2) the part of the visual display that is in the subject's foveal view
> Newell and Simon, 1972, p.

The feeling of visual richness, I want to suggest, is thus a bit like the feeling of 'knowing a lot about pulp detective novels'. It is not that all that arcane knowledge is there all at once, actively co-present in conscious awareness. Rather, what is present is a kind of meta-knowledge: the knowledge that you can retrieve just about any relevant bit of all that information as and when required, and deploy it in the service of your current conscious goals. Our conscious experience of visual richness, if this is at all on track, is an experience of a kind of *problem-solving poise.*

All this, it seems to me, is correct and important. But the rather specific version of the skill-theory presented by O'Regan and Noe (In Press-2001) is more radical in at least two respects. First, because it depicts the skill theory as a direct dissolution of the 'hard problem' of visual qualia. Second, because it endorses (what I suspect to be) an overly motoric view of the wellsprings of conscious visual awareness.

My goal in the remainder of the paper will be to flesh out these worries, and to offer a weakened (regarding the 'hard problem') and amended (regarding the role of motor action) version of a skill-theoretic account

V.   Sensorimotor Chauvinism and The Hard Problem

O'Regan and Noe offer an unusually clear, and refreshingly ambitious story. The 'hard problem' of explaining visual qualia (what it is like to see red, why it is like anything at all to see red, etc) is, they suggest, simply unable to arise. And the explanatory gap thus feared between scientific accounts and the understanding of qualitative consciousness is no gap at all. The trouble arises, they suggest, only if we falsely believe that visual qualia are properties of experiential states. And this is (it is claimed) a theoretical mis-step since:

> Experiences…are not states. They are ways of acting. They are things we do…there are, in this sense at least, no (visual) qualia. Qualia are an illusion and the explanatory gap is no real gap at all.
> O'Regan and Noe (In Press-2001) p.25

Dispelling the Grand Illusion illusion, it now seems, requires us to embrace an even greater oddity: the idea that qualia, properly speaking, do not exist! To sweeten the medicine, the authors use the familiar (ok, so mine's a Ford) example of driving a Porsche. There is, they admit, 'something that it is like to drive a Porsche' But this "something it is like" does not consist in the occurrence of a special kind of internal representation (the kind supposedly accompanied by qualia). Rather it consists in "one's comfortable exercise of one's knowledge of sensorimotor contingencies governing the behaviour of the car" (Op cit., p.25). The driver knows, that is to say, how the car will corner, accelerate, and respond to braking, and much more besides. Most of this knowledge is non-propositional, more in the

realm of skilled know-how than reflective awareness. But knowing what it is like to drive a Porsche just is, the authors argue, having a bunch of such know-how. Similarly, knowing what it is like to see a red cube is simply knowing how the image of the cube will distort and alter as you move your eyes, how uneven illumination will affect the inputs, etc etc. In all these cases it is the fact that our knowledge is implicit, knowing-how not knowing-that, that makes it seem as if there is something 'ineffable' going on (op cit. p. 26).

But this is not convincing. Consider a fairly simple Ping-Pong playing robot[ix]: a descendent, perhaps, of the fairly successful prototype described in Andersson (1988). The robot uses multiple cameras, and it has an arm and a paddle. A modest on-line planning system plots initial paddle-to-ball trajectories, but this is soon improved during play and practice, as the system learns to use simpler visual cues to streamline and tune its behaviour. The robot, let us suppose, develops (courtesy of a neural network controller) a body of implicit knowledge of the relevant sensorimotor contingencies. Finally, it is able to deploy this knowledge in the service of some simple goals, such as the goal of winning, but not by more than 3 points. At this moment, as far as I can tell, all of O'Regan and Noe's conditions have been met:

> 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 sensorimotor contingencies, it must make use of this exercise for the purposes of thought and planning.
> O'Regan and Noe (In Press-2001) p.7

Assuming, then, that the term 'thought' is not here begging the question (by meaning something like 'experience-accompanied reasoning'), the Ping-Pong robot is a locus of qualitative visual experience. But while a few philosophers might take a deep breath and agree, I suggest that the attribution of qualitative consciousness is fairly obviously out of place here. O'Regan and Noe (2001) appeared to bite the bullet, by endorsing a kind of continuum view of qualitative consciousness, and allowing that this kind of robot would

indeed have some of it . In a footnote to the paper (note 7) they write 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 they are also aware" (op cit, note 7, page 46)

The question, I think, is whether we should *at this very early stage in the investigation of qualitative consciousness* simply give up on the main intuitions that currently demarcate the very target of our theorising (intuitions such as: we have it, the Ping Pong Robot doesn't, and we aren't sure about a lot of animals). In my view, the price of giving up these intuitions so soon is that we will never know when (or if) we have explained what we set out to. The pay-off (a very neat theory) is surely not worth this cost. Moreover we know, from our own experience, that visual information can guide apparently goal based activity while we are quite unaware of it doing so. Back in the Porsche, we may make a successful turn, to head for home, while fully engaged in some other task. Knowing just how much we ourselves can achieve with non-conscious sub-systems at the wheel, we are rightly suspicious of attributing too much too soon to the Ping Pong Robots of this world.

In response to the Ping-Pong playing robot counter-example, O'Regan and Noe suggest that the robot described is " far too simple to be a plausible candidate for perceptual consciousness of the kind usually attributed to animals or humans" (O'Regan and Noe (In-Press B 2002) p.4). This simplicity is said to consist both in a lack of advanced sensorimotor skills and in the absence of a thick background of intentions, thoughts, concepts and language (op cit). Nonetheless, it still seems to me that the robot described meets the letter of the requirements laid out, and should (on their official account) be granted some small degree of conscious visual awareness- even if not 'of the kind' (though this is a somewhat vague and elusive notion) attributed to more complex beings

In addition, as Mark Rowlands[x] has argued, even the general form of the attempted dissolution of the hard problem is suspect. First, since the hard problem arises for all varieties of phenomenal experience, it seems fair to ask how well the skill-based account will fare with the others? And for cases such as the feeling of depression, elation, pain etc it is not at all clear how it will work. Second, from the fact that a certain theoretical gloss on the hard problem is rejected, we cannot conclude that the problem itself has gone away. Thus we may agree that it is misleading to think of the hard problem as the problem of how certain internal representations come to generate qualitative experience. But it would not follow that the question, How is such experience possible? is somehow mis-posed. Taking the full skill-based story on board, the question can still be asked: Why is it like *anything at all* to see red, to drive a Porsche etc?

The hard problem really has two components which need to be kept distinct. The first is: Why is such-and-such an experience like this rather than like that (why does Marmite taste like this and not like something else?). The second is: Why is it like anything at all? Many theories that get a purchase on the former fail to illuminate the latter, and the skill-theory belongs in this camp. The pattern of sensorimotor contingencies may help explain why experiences have the contents they do, but not why it is like anything at all to have them. (Conversely, accounts such as Clark (2000), which try to address the latter question, often fail to say anything about the former. We learn why it seems like something, but not why it seems the way it does!).

From here on, I shall understand the skill-theory as an attempt to shed light on why certain experiences seem the way they do, rather than why they seem like anything at all. Even thus understood, the O'Regan and Noe proposal faces a rather important challenge. For the way the story is developed, it runs the risk (or so it seems to me) of a certain kind of *over-sensitivity to low-level motoric variation.* This kind of over-sensitivity I shall label 'sensorimotor chauvinism'. Here's what I have in mind.

It is an implication of the way O'Regan and Noe develop the skill theory that my conscious visual experience depends very very sensitively upon my implicit knowledge of a very specific set of sensorimotor contingencies, including those that they term 'apparatus-related' i.e. relating to the body and sensory apparatus itself. Now certainly, they want to allow that what is broadly speaking *visual* experience could indeed be supported by many different kinds of sensing device, including TVSS arrays etc. It is the structure of the rules of sensorimotor contingency that matters, not the stuff. Nonetheless, it is equally clear that very small differences in the body and sensory apparatus will make a substantial difference to the precise set of sensorimotor contingencies that are implicitly known. And indeed, the authors are at pains to stress the importance of, for example, the precise way the sensory stimulation on the retina shifts and changes as we move our eyes (op cit. p.3), as helping to fix the pattern of sensorimotor contingencies. But of course it is (on their account) this very pattern that in turn determines the nature and content of our conscious visual experience.

The suspicion I want to voice, then, is that this may make the contents of my visual experience *too* sensitive to the very precise, low-level details of sensory pick-up and apparatus. Suppose, for example, that my eyes saccade fractionally faster than yours. This will change the pattern of sensorimotor contingencies. But why should we believe that *every* such change in this pattern will yield a change, however minute, in the nature and contents of my conscious visual awareness? If we don't believe this, then we will want to know what makes it the case that *some* changes in patterns of sensorimotor contingency impact conscious visual experience and some don't. In sum, O'Regan and Noe must either accept that *every difference makes a difference*, or they owe us an account of which one's matter and why.

In response to this charge, O'Regan and Noe (In Press B 2001) embrace the idea that every difference makes a difference. Indeed, they embrace the strongest form of this idea, saying that their view:

Allows for the judgement that creatures with radically different kinds of physical make-up can enjoy experience which is, to an important degree, the same in content and quality. But it also allows for the possibility *(indeed the necessity) that where there are physical differences, there are also qualitative differences.*
O'Regan and Noe (2001, p.4. My emphasis)

It is this latter consequence which I shall reject. The question what differences make a difference should, I believe, be an open empirical question. It should not be foreclosed by an overly enthusiastic development of the skill- theory. For skill theory, as I have argued elsewhere (Clark (1999), (In Press)) , has the resources to allow for many different kinds of way in which cognition may be "action-oriented" and some of these leave plenty of room for loosening the ties between the full gamut of physical apparatus and the contents and character of conscious experience.

Indeed, the resolution of this conundrum is hinted at by O'Regan and Noe's own (very proper) insistence on the importance, for conscious vision, of linking currently exercised mastery of sensorimotor contingencies to planning, deliberation and intentional action. Mastery of the laws of sensorimotor contingency, O'Regan and Noe insist, must be 'exercised for the purposes of thought and planning' (op cit. p.7). But this very role, I shall next suggest, may act as a kind of filter on the type and level of detail (of mastery of sensorimotor contingencies) that matters for the determination of the contents of conscious visual experience.

VI.   Reason, Action and Experience

Here's where we seem to be. We began with the idea that our visual experience may not be all it seems: that we may be misled into thinking we see more, and are sensitive to more changes, than we actually are. Careful examination suggests, however, that what is really at fault is a certain theoretical model of that in which conscious seeing consists. If conscious seeing were forced to consist in the

internal tokening, moment-by-moment, of a constantly updated model of the scene, then we would indeed be wildly misled by our experiences of visual richness. For our persisting internal representations, such as they are, appear to be sparse and high-level, supplemented by more detailed information retrieved at the last possible moment, and kept only for the duration of the appropriate element of the task.

The skill-theoretic approach offers an alternative theoretical model, relative to which our error is much less dramatic. By highlighting the way temporally extended information-seeking activity actually *constitutes* successful contact with a rich and detailed external scene, skill theory dispels the miasma of Grand Illusion. Our visual experience reflects our successful engagement with the richness of the scene. O'Regan and Noe's marvellously detailed development of the skill-theoretic line threatens, however, to tie conscious visual experience too closely to the precise details of the low-level sensorimotor routines by means of which this engagement proceeds. Yet the resources are all there to support a slightly different kind of skill-theoretic story. For O'Regan and Noe also insist, importantly, on the profound connection between conscious visual experience and intentional action: the kinds of action we would describe as 'deliberate' and as emanating from the conscious endorsement of reasons and plans. Proper attention to this dimension suggests a slightly different way to develop and deploy the skill-theoretic intuitions.

We can creep up on this by highlighting a second way in which we might perhaps be accused of misunderstanding the nature and role of our own conscious visual experience. This is by making what I call (Clark, In Press) the Assumption of Experience-Based Control:

> **(Assumption of Experience-Based Control. EBC.)**
>
> The details of our conscious visual experience are what guide fine-tuned motor activity in the here-and-now.

We often relax this assumption when reflecting upon, for example, our experiences of playing sports etc. At such moments we realise that there is really no way our conscious visual experience is fine-tuning our actions. But we seem to believe, for the most part, that our conscious seeings are usually guiding and controlling our visually-based activities. To see ourselves aright, then, it is important to be very clear about the precise sense of control and guidance that is most likely actually at work.

Taken at face value, the assumption of experience based control is increasingly suspect. Thus consider Milner and Goodale's provocative claim that "what we think we 'see' is not what guides our actions" (Milner and Goodale 1995 p.177). The idea, which will be familiar to many readers, is that online visually guided action is supported by neural resources that are fundamentally distinct from, and at least quasi-independent of, those that support conscious visual experience, off-line imagistic reasoning, and visual categorisation and planning. More specifically, the claim is that the human cognitive architecture includes two fairly distinct 'visual brains'. One, the more ancient, is specialised for the visually-based control of here-and-now fine motor action. The other, more recent, is dedicated to the explicit-knowledge-and-memory based selection of deliberate and planned actions (what Milner and Goodale (1998, p.4) nicely describe as 'insight, hindsight and foresight about the visual world'). The former is then identified with the dorsal visual-processing stream leading to the posterior parietal lobule, and the latter with the ventral stream projecting to inferotemporal cortex.

Computationally, some such division of labour makes good sense. The fine-grained control of action (the precise details of the visually-guided reach for the coffee-cup etc) requires rapidly processed, constantly updated, egocentrically specified information about form, orientation, distance etc. Conceptual thought (the identification of objects and the selection of deliberate actions) requires the identification of objects and situations according to category and significance, quite irrespective of the precise details of retinal image

size etc. Achieving a computationally efficient coding for either of these pretty much precludes the use of that very same coding for the other. In each case, as Milner and Goodale note, we need to extract, filter, and throw away different aspects of the signal, and to perform very different kinds of operation and transformation.

Concrete evidence in support of the dual visual systems view comes in three main forms. First, single cell recordings that show the different response characteristics of cells in the two streams. For example, PP (posterior parietal) neurons that respond maximally to combinations of visual cues and motor actions, and IT (inferotemporal) neurons that prefer complex object-centred features independently of location in egocentric space (see e.g. Milner and Goodale (1995) p.63).

Second, there are the various pathologies. The most famous example is DF, a ventrally-compromised patient who claims she has no conscious visual experience of the shape and orientation of objects but who can nonetheless perform quite fluent motor actions (such as pre-orienting and posting a letter through a visually presented slot). Importantly, DF fails to perform well if a time delay is introduced between presentation of the slot and selection (with the slot now out of sight) of an orientation. This presumably shifts the burden from the intact, putatively non-conscious dorsal stream to the impaired ventral resource dedicated to memory, planning and deliberate action selection. Optic ataxics, conversely, are dorsally impaired and claim to see the objects perfectly well despite being unable to engage them by fluent behaviours. These patients are actually helped by the introduction of a short time delay.

Third, there is some (controversial) evidence from normal subjects. Certain visual illusions[xi], for example, seem to affect our conscious perceptions without impairing our ongoing visuomotor motions. In these cases, Milner and Goodale suggest, the non-conscious dorsal stream controls the fine-tuned motions and is immune to the illusion, which arises due to processing idiosyncrasies in the ventral stream.

(And in these cases, likewise, introduction of a time-delay blocks the accurate performance).

Milner and Goodale end their account with a model of how the two visual brains interact. The interaction, they suggest, occurs precisely at the level of intentional agency. Conscious visual experience can select the targets of actions, and the types of action to be performed. For example, conscious vision is used to select the red cup on the left and to decide on a grip appropriate to throwing rather than drinking. But it is then left to the non-conscious dorsal stream to work out how to implement these plans and ideas.

Milner and Goodale, it is reasonable to suspect (and see Clark (1999) (In Press) for some details), overplay the extent to which the two streams work in near-isolation. For example, Pascual-Leone and Walsh (2001), in an elegant application of TMS (transcranial magnetic stimulation) show that feedback from high-to-low level visual areas (from V5/MT to V1 and V2) is necessary for certain kinds of conscious visual perception. This opens up the intriguing possibility that upstream activity of many kinds could directly modify conscious visual awareness by altering activity at the common gateway to both streams. And there are, without doubt, many complex and iterated interactions which compromise the isolationist integrity of the two streams. Moreover, there is a convincing case to be made that the degree of stream-independence, and (conversely) the nature and extent of stream interaction, is both task and attention dependent (see Brennar and Smeets (1996), Jeannerod (1997) Decety and Grezes (1999), Rensink (2000), Carey (2001) and discussion in Clark (In Press)).

A weakened version of the dual visual systems hypothesis, however, enjoys widespread support (e.g. Jeannerod (1997), Decety and Grezes (1999)). Such accounts accept the task-variability of the inter-stream relationship, and leave room for complex feedback modulated interactions, but they preserve the essential insight, which is that substantial amounts of fine-action-guiding visual processing are often carried out independently of the processing underlying conscious

visual awareness. Such accounts accept that when we keep looking while performing a task, we are indeed feeding two sets of (partially interacting) processes: one more ancient, concerned with fine visuomotor action in the here-and-now, the other more recent, and geared towards reasoning and conscious awareness. It is the latter system, geared towards spotting the meaningful in a way fit for reasoned action-selection, that is also most closely associated with semantic and episodic memory systems.

Consider, to take just one more example, a series of experiments in which subjects were required to both visually track and manually point out a visually presented target. This target, however, was sometimes suddenly (unexpectedly) slightly displaced after the original presentation. Bridgeman et al (1979) showed that subjects would accommodate this displacement (as evidenced by accurate saccades and pointing) whilst remaining quite unaware that the target had moved. Moreover, in those cases where the displacement was large enough to attract attention and hence to enter conscious awareness, the on-line adjustments were much less fluid and less successful (for a rehearsal, see Milner and Goodale, 1995, 161).

To round this story off, Wong and Mack (1981) showed that subjects who automatically and unconsciously accommodate the smaller displacements will, if subsequently asked to point to the remembered location of the (now-removed) target, actually point to the original (non-displaced) location. Similar results have been obtained for grasping motions directed at present versus remembered visually displayed objects (see Milner and Goodale 1995, 170-173). Memory-driven responses thus seem to be tied to the contents of conscious visual experience, while on-line object-engaging performance is driven by a distinct and more sensitive resource.

The Assumption of Experience-Based Control thus needs to handled with extreme care. Our conscious visual experiences certainly impact our choices of actions. But they do not do so by virtue of providing the visual information that is itself used for the fine-tuned control of

movement. Clark (In Press) suggests that the EBC should thus be replaced by something like this:

Hypothesis of Experience-based Selection (EBS)

Conscious visual experience presents the world to a subject in a form appropriate for the reason-and-memory based selection of actions.

In all the cases we have discussed, the alignment of certain memory-systems with conscious visual experience looks robust and significant. This simple fact leads to the final idea that I want to consider. It is the idea that, contrary to the most radical versions of the skill theory:

The key to connecting consciousness with action might involve memory systems rather than motor systems
Prinz (2000) p.252

Prinz' speculation is that the evolution of new episodic and working memory systems fundamentally altered- for certain organisms-the relation between perception and action. Phylogenetically more ancient structures could already support the rapid, input-driven selection of innate and learnt motor responses, and could initiate whole cycles of environmental probing in which sensing and acting are deeply interanimated. But in some animals new working memory systems began to support the retention and off-line manipulation of perceptual information. Episodic memory systems allowed them "to encode particular perceptual events in a long-term store and to access those events on future occasions" (op cit. p.253). These explicit memories could be called up even when the circumstances to which they were initially keyed were no longer present, and put into contact with systems for planning (real planning, in which multiple stages of action are considered, chained together, and assessed) and reasoned action-selection. The emergence of these new reason-and-memory-based systems marked, Prinz speculates, the emergence of consciousness itself. In a similar vein[xii] Hardcastle (1995), following an extensive

review of the neuroscientific and psychological literature, suggests that:

> Conscious perceptions and thoughts just are activations in SE [semantic, 'explicit controlled access'] memory
>
> Hardcastle 1995 p.101

One way or another, then, the links between special kinds of memory systems and conscious experience seem strong. Conscious experience is, above all, the base for reasoned, deliberate action selection. And this requires deep and abiding links with the special memory systems that mediate between sensory input and action.

I should add one important caveat. In speaking of the importance of explicit memory structures, I make no commitment to any specific story about encoding or storage. In particular, it seems highly unlikely that such encodings are in any interesting sense propositional. Instead, the stored information is most likely geared quite tightly to the kinds of action we may need to select, and to the environmental resources upon which we may reasonably rely.

With this in mind, let us finally revisit the rather strong form of the skill-theory as advanced in O'Regan and Noe (2001) . Here, knowledge of a specific set of potential movements and their results is said to constitute a given visual perception (Op Cit p.13). The general idea is that knowledge of the laws of sensorimotor contingency actually constitutes the way the brain codes for visual features and attributes. This idea is economical, elegant and attractive. But in the specific case of *conscious* visual perception, the work on the dual visual systems hypothesis suggests an alternative unpacking. For what matters, as far as conscious seeing is concerned, is that the object/event is 'one of those' (ie falls into such-and-such a class or category) and that a certain range of actions (not movements, but actions such as grasping-to-throw, grasping-to-drink, etc) is potentially available. Both 'visual brains', I am suggesting, represent by activating implicit knowledge of some set of possible actions and results. In the case of the 'visuomotor brain' these are indeed pitched

at the kind of level O'Regan and Noe seem to favour: they will concern e.g. the anticipated distortions of the retinal image in response to certain head and eye motions etc. But in the case of the 'conscious visual brain' they are more likely to concern types of action and their effects as applied to types of objects eg the way 'throwing the cup at the wall' gives way to 'smashed cup on the floor' and so on. These kinds of sparse, high-level understanding are, of course, precisely the kinds of understanding that do seem to underpin our conscious visual experience, as the various change blindness results (and the card tricks) help to show.

One way to dramatise this idea is to exploit the idea (Goodale (1998)) that conscious seeing acts in a way somewhat reminiscent of the interaction between a human operator and a smart teleassistance device. The operator decides on the target and action-type (for example "pick up the blue rock on the far left") and the robot uses its own sensing and acting routines to do the rest. Knowledge of our capacity to *engage* such routines may, on the present account, be essential to the content of the experience, even if the routines themselves employ sensory inputs in a very different, and largely independent, way.

O'Regan and Noe claimed, recall, that

> For a creature 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*
>
> Op Cit p 7 My emphasis

But what exactly does this mean? Imagine again a tele-assistance set-up in which the distant robot has implicit mastery of the SMC for, say, reaching and grabbing.This will only *matter*, as far as the conscious controller is concerned, in a functional way: the controller needs to know what the robot can and can't do (it can't fly, it can reach, it can grab gently or harder, etc). *The SMC knowledge that the*

robot depends upon could be quite different, in detail, as long as the broad functionality was the same.

What matters for conscious vision, on this alternative model, is that the visually seen object is recognised as belonging to some class, and as affording certain types of action. The bodies of know-how that count here concern objects and events at this level of description. Insofar as the lower-level SMC's matter here, they do so non-constitutively. Sameness of visual experience thus depends on sameness of what might be called 'intentional role' rather than sameness of all the SMCs.

At best, then, there is an unclarity hereabouts in O'Regan and Noe's account. The skill-theoretic response to the Grand Illusion story is, I think, the right one. But it needs to take two forms to deal with the full gamut of ways the brain uses visual information. One of those ways is geared to the fine control of here-and-now visuomotor action, and the relevant laws of sensorimotor contingency here do indeed concern the very precise- and fully apparatus dependent- ways that the eyes are stimulated in response to various kinds of motion and probing. The other is geared to the selection of actions (not motions) and to planning and reasoning, and the relevant implicit knowledge here takes a different and more 'meaningful' form: it is knowledge of what we can do and achieve on the basis of current visual input, knowledge of a space of actions and results, rather than of a space of movements and subsequent inputs[xiii].

I would be the last to downplay the significance, in human cognition, of tightly coupled, embodied, embedded sensorimotor loops (see Clark 1997). But these loops, in the case of humans and (I expect) other higher animals are now themselves intertwined with new circuitry geared towards knowing, recall and reasoning. Understanding both the intimacies *and the estrangements* that obtain between these recently coiled cognitive serpents is, I suggest, one of the most important tasks facing contemporary cognitive science.

VII.   Conclusions: Seeing a World Fit to Think In.

The skill theory, as developed by O'Regan and Noe, ties conscious visual experience rather too closely (I have argued) to the precise details of our sensory engagements with the world. If (say) my eyes saccade just a little faster than yours, this may have no impact upon the qualitative nature of my visual experience. For conscious vision is geared to presenting the world for reason and for quite high-level action selection. This requires converging on-the-spot visual input (gathered just-in-time, and as dictated by the task and the allocation of attention) with stored memories and expectations. What matters for visual consciousness is thus (I suggest) at best a select subset of the information O'Regan and Noe highlight. The full detail of the sensorimotor contingencies that characterise my visual contact with specific objects and events is unlikely to matter. What will matter are whatever (perhaps quite high-level) aspects of those sensorimotor contingencies prove most useful for reason, recognition and planning. If this is correct, it is a mistake to tie visual experience too tightly to the invariants that guide and characterise visuomotor action.

Where the skill theory scores, however, is in recognising that conscious perceptual experience need not (and should not) be identified with a single time-slice of an environmentally isolated system. Instead, we need to consider the way temporally extended sequences of exploratory actions, and our knowledge of the availability and likely deliverances of such exploratory routines, may actually help constitute the contents of my perceptual experience. And this, in turn, requires recognising the way the external scene may itself feature as a kind of temporary memory resource, able to be accessed and deployed as and when the task requires.

As for the Grand Illusion, that really was a trick of the light. For once we take all this into account, our visual experience is not itself misleading. The scene before us is indeed rich in colour, depth and detail, just as we take it to be. And we have access to this depth and detail as easily as we have access to facts stored in biological long-term memory. It is just that in the case of the visual scene, retrieval is via visual saccade and exploratory action. Our daily experience only

becomes misleading in the context of a host of unwise theoretical moves and commitments: commitments concerning the precise role of internal representations in supporting visual experience, as well as our pervasive neglect of the cognitive role of temporally extended processes and active exploration.

A full account of conscious seeing cannot, however, stop there. For the world is seen, via these exploratory engagements, in a way that continuously converges selective input sampling with stored knowledge, memories and expectations. The contents of conscious visual experience emerge at this complex intersection. What we consciously see is a world tailor-made for thought, reason and planning. That's why we can intentionally act in the very world we experience.

References

Andersson, R.L (1988) *A Robot Ping-Pong Player* (MIT Press, Cambridge, MA)

Ballard, D. (1991). Animate Vision. *Artificial Intelligence*, 48, 57-86.

Ballard, D., Hayhoe, M., Pook, P. and Rao, R. (1997). Deictic Codes for the Embodiment of Cognition. *Behavioral and Brain Sciences*, 20, 4.

Brenner, E., and Smeets, J. (1996): "Size Illusions Influence How We Read But Not How We Grasp An Object". *Experimental Brain Research*, 111:473-476.

Bridgeman, B., Lewis, S., Heit,G., and Nagle, M. (1979) "Relation between cognitive and motor-oriented systems of visual position perception" *Journal of Experimental Psychology (Human Perception*) 5:692-700

Carey, D. (2001) "Do Action Systems Resist Visual Illusions*?" Trends in Cognitive Sciences* 5:3: 109-113

Churchland, P.S., Ramachandran, V. and Sejnowski, T. (1994). A Critique of Pure Vision. In C. Koch and J. Davis (eds.), *Large-Scale Neuronal Theories of the Brain*, Cambridge, MA: MIT Press.

Clark, A. (1993). *Associative Engines: Connectionism, Concepts and Representational Change*, Cambridge, MA: MIT Press.

Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again*. Cambridge, MA: MIT Press.

Clark, A (1999) "Visual Awareness and Visuomotor Action". *Journal of Consciousness Studies*. 6:11-12: p.1-18

Clark, A (2000) "A Case Where Access Implies Qualia?" *Analysis* 60:1: 30-38

Clark, A (In Press) "Visual Experience and Motor Action: Are the Bonds Too Tight?" *Philosophical Review*

Clark, A. and Chalmers, D. (1998). The Extended Mind, *Analysis,* 58: 7-19.

Clark, A and Toribio, J (In Press-2001) " Sensorimotor Chauvinism? Commentary on O'Regan and Noe "A Sensorimotor Account of Vision and Visual Consciousness" " in *Behavioral and Brain Sciences* 24:5.

Decety, J. and Grezes, J. (1999) "Neural Mechanisms Subserving the Perception of Human Actions". *Trends in Cognitive Sciences, 3 :5: 172-178.*

Dennett, D (1991) *Consciousness Explained* New York, Little Brown

Feldman, J.A. (1985) "Four Frames Suffice: A Provisional Model of Vision and Space" *Behavioral And Brain Sciences* 8:265-289

Goodale, M. 1998 "Where Does Vision End and Action Begin?" *Current Biology,* 489-R491.

Hardcaste, V (1995) *Locating Consciousness* (John Benjamins, Amsterdam)

Jeannerod, M. (1997) *The Cognitive Neuroscience of Action.* Oxford: Blackwell.

Kirsh, D (1991) "When is information explicitly represented?" in P. Hanson (ed) *Information, Thought And Content* (UBC Press, Vancouver)

Mack, A and Rock, I (1998*) Inattentional Blindness* (MIT Press, Camb. Ma)

MacKay, D (1967) "Ways of looking at perception" inW. Wathen-Dunn (ed) *Models For The Perception Of Speech And Visual Form* (MIT Press, Camb. MA) 25-43

McConkie, G. W. (1990). "Where vision and cognition meet". Paper presented at the H.F.S.P. Workshop on Object and Scene Perception, Leuven, Belgium.

Milner, D. and Goodale, M. (1995) *The Visual Brain in Action.* Oxford: Oxford University Press.

Milner, D. and Goodale, M. (1998) "The visual brain in action (precis)". *Psyche*, 4 (12), October 1998.

Neisser, U (1979) "The control of information pickup in selective looking" in A.Pick (ed) *Perception And Its Development* (Erlbaum:N.J.) 201-219

Newell, A and Simon, H (1972*) Human Problem Solving* (Prentice Hall, N.J.)

Noe, A., Pessoa, L, Thompson, E. (2000) Beyond the grand illusion: what change blindness really teaches us about vision. *Visual Cognition* 7: 93-106.

Noe, A. (2001) Experience and the active mind. *Synthes*E 129:41-60.

Noë, A. and O'Regan, J. K. (2000). Perception, attention and the grand illusion. *Psyche*, 6(15).

O'Regan, J.K. (1990) "Eye movements and reading" in Kowler, E (ed) *Eye Movements And Their Role In Visual And Cognitive Processes* (Elsevier, Amsterdam)

O'Regan, J.K. (1992) "Solving the "real" mysteries of visual perception: the world as an outside memory" *Canadian Journal Of Psychology* 46:3:461-488

O'Regan, J.K., and Noe, A In Press, 2001 "A Sensorimotor Account of Vision and Visual Consciousness" *Behavioral and Brain Sciences* 24:5.

O'Regan, J.K., and Noe, A In Press B 2001 "Authors' Response: Acting Out Our Sensory Experience" *Behavioral and Brain Sciences* 24:5.

Pascual-Leone, A and Walsh, V (2001) "Fast backprojections from the motion to the primary visual area necessary for visual awareness" *Science* 292:510-512

Pessoa, L., Thompson, E., and Noë, A. (1998). Finding out about filling in: A guide to perceptual completion for visual science and the philosophy of perception. *Behavioral and Brain Sciences*, 21(6):723-802.

Prinz, J (2000) "The Ins and Outs of Consciousness". Brain and Mind 1 (2) :245-256

Rensink, R (2000) "Seeing, Sensing and Scrutinizing" *Vision Research* 40: 1469-1487

Rowlands, M (ms) "Two Dogmas of Consciousness" available at: www.ucc.ie/ucc/depts/phil/

Simons, D (2000) "Attentional capture and inattentional blindness" " *Trends in Cognitive Sciences*,: 4:147-155

Simons, D and Levin, D (1997) "Change Blindness" *Trends in Cognitive Sciences*, 1: 7: 261-267

Simons, D and Chabris, C "Gorillas in our midst: sustained inattentional blindness for dynamic events" *Perception* 28 1059-1074

Verghese, P and Pelli, D (1992) "The information capacity of visual attention" *Vision Research* 32: 983-995

Wolfe, J (1999) "Inattentional Amnesia" in V. Coltheart (ed*) Fleeting Memories* (MIT Press, Camb. MA)

Wong,E., and Mack,A. 1981 "Saccadic programming and perceived location" *Acta Psychol*., 48: 123-131

Yarbus, A. (1967). *Eye movements and vision*. New York: Plenum Press.

---

[i] http://members.tripod.com/~andybauch/magic.html. Or just feed 'amazing card trick' to a search engine such as Google.

[ii] Recall Dennett's 'many Marilyns' example, as described in Dennett 1991

[iii] To my knowledge, the phrase Grand Illusion was first used in Noe, Pessoa and Thompson (2000), which was a critique of the idea that visual experience involved any such illusion.

[iv] Most of the experiments descibed below can be viewed on the web. Try http://nivea.psycho.univ-paris5.fr, http://coglab.wjh.harvard.edu, http://www.cbr.com/~rensink

[v] Yarbus (1967).

[vi] McConkie, G. (1990), O'Regan (1990). And see discussion in Churchland et al (1994)

[vii] McConkie (1990), O'Regan (1992).

[viii] The original experiment was done by Neisser (1979). Recent versions, including the opaque case, are due to Simons and Chabris (1999).

[ix] This example is from Clark and Toribio (In-Press, 2001)

[x] See the draft paper "Two Dogmas of Consciousness" on his web page at: www.ucc.ie/ucc/depts/phil/

[xi] E.g. the Tichener Circles illusion discussed in Milner and Goodale (1995) Ch. 6 – see Clark (In Press) for an extended discussion of this case

[xii] The accounts are by no means identical. Prinz emphasizes akind of informational poise at the gateways to the memory systems, whereas Hardcastle emphasizes activity in the memory systems themselves. See especially Prinz (2000) p.255

[xiii] For a little more on this, see Clark (1999) section 3 "two ways to be action-oriented".