Mechanisms of Recognition: Change-blindness and the Perception of Pictures
By Ron Gallagher

We see with the brain, not with the eyes. Paul Bach-y-Rita

How does our perception of the brown smudges in Constable’s Ottawa study of Salisbury Cathedral (fig. 1) relate to how we see cows in real life? If Flint Schier is to be believed, looking at that brown smudge is triggering something in us that is triggered by an encounter with the real thing. This seems unfeasible. Even when we look at the Wivenhoe Park cows (fig.2 -Friesians and Herefords, I think) it is still hard to say that this is much like seeing real cows.

The working hypothesis of this thesis is that the techniques of artists and photographers mobilize the same mechanisms of recognition which we use to see objects and scenes in the world. Schier has proved a reliable guide so far, and his “two-hypotheses” thesis requires that
we entertain the possibility, however remote, that our visual system at some basic level takes these paint marks to be cows. In short, I need to show that seeing a picture of a cow is far more like seeing a cow in real life than you might suppose. I propose to demonstrate this by relating aspects of how we see pictures to aspects of how we see things in the world. Following Michael Podro, I will show how artists “sustain recognition” by mobilizing the mechanisms of visual cognition. My first example shows how pictorial composition mobilizes recognitional abilities.

**Composition, Change Blindness and Visual Metacognition**

David Wooding’s experiment at the National Gallery’s “Telling Time” exhibition in 2000-2001 tracked the eye movements of over five thousand gallery attendees as they looked at paintings. The results show that in the majority of cases we only look carefully at parts of the picture towards which the artist’s composition directs our eye. One of the paintings in the study was Veronese’s *Christ addressing a Kneeling Woman* (fig. 3) and we can see that viewers’ eyes have been strongly drawn to the upper torso of the two main figures, to the exclusion of the other parts of the painting.

![Figure 3](image)

Veronese has directed the viewers’ attention through colour and brightness contrast, and by connecting the gazes of the main characters. This compositional technique is augmented by the alignment of heads and the gaze of the other characters who are looking towards Christ and the kneeling woman.
What Wooding’s experiment reveals is that contrary to Gombrich, the symmetry of composition does not work against representation of the subject, but enhances it by capitalising on mechanisms of visual cognition. We are drawn to look at the two main characters because Veronese is manipulating our visual mechanism using gaze, colour light and composition. Podro takes issue with Gombrich treating “convincing representation and the pursuit of pictorial order as reciprocally curtailing” and argues:

these demands may sustain each other: symmetries or correspondences are ways in which we see one form on analogy with another and project one aspect onto another; this is integral to how recognition is sustained and developed.¹

Podro examines a similar technique of alignment, gaze and symmetry being used in Ghiberti’s panel of the Adoration and Leonardo’s Adoration. He further argues that, not only do deliberate techniques of the artist sustain recognition of the subject, but seemingly accidental features of a drawing or paintings, such as unfinished edges, roughly sketched lines and the texture of the paint, are recruited to the depiction of the subject.

Figure 4 Leonardo Da Vinci, Head of a girl, c. 1483, Silverpoint and white highlights on prepared paper, 181 x 159 mm, Biblioteca Reale, Turin

If we apply this line of thinking to the Leonardo drawing Head of a Girl, which we discussed in the last chapter (fig. 4), we can imagine that if we used Wooding’s eye-tracking technique to reveal which parts of the picture we were most fixated-on, it would almost certainly be the girl’s eyes. The rest of the picture would hardly get a glance. Thus, one reason Leonardo hasn’t rendered the rest of the picture is because he knows that, given his composition, the viewer simply wouldn’t attend to the hair and the shoulders even if he had drawn them.

¹ Michael Podro, Depiction (Yale University Press, 1998), 8-9.
This explanation seems at odds with Schier’s speculation that it is “frame knowledge” that enables us to discount these accidents of technique. In fact, the two explanations are compatible – both claim that our visual system effectively ignores or throws away a good deal of the visual information that presents to our field of view. This is confirmed by the research on change blindness and inattention blindness that has been conducted over the last few years. All the evidence from these studies indicates that we notice and remember very little from a visual scene. Furthermore, we vastly overestimate our own and other people’s abilities to attend to and recall what we see. Rather curiously, this phenomenon was predicted by Daniel Dennett in 1991.² In a review of change blindness research, Laura Spinney summarises its impact on vision research:

Until the last decade, vision researchers thought that seeing really meant making pictures in the brain. By building detailed internal representations of the world, and comparing them over time, we would be able to pick out anything that changed. Then in 1991, in his book *Consciousness Explained*, the philosopher Daniel Dennett made the then controversial claim that our brains hold only a few salient details about the world— and that this is the reason we are able to function at all.

We don't store elaborate pictures in short-term memory, Dennett said, because it isn't necessary and would take up valuable computing power. Rather, we log what has changed and assume the rest has stayed the same. Of course, this is bound to mean that we miss a few details. Experimenters had already shown that we may ignore items in the visual field if they appear not to be significant—a repeated word or line on a page of text, for instance. But nobody, not even Dennett, realised quite how little we really do "see".

Just a year later, at a conference on perception in Vancouver, British Columbia, John Grimes of the University of Illinois caused a stir when he described how people shown computer-generated pictures of natural scenes were blind to changes that were made during an eye movement. Dennett was delighted. "I wish in retrospect that I'd been more daring, since the effects are stronger than I claimed," he says.³

Grimes’ experiments tested whether subjects noticed changes in a scene introduced between saccades. In his 1996 paper, “On the Failure to Detect Changes in Scenes across Saccades” Grimes concludes that very little information is carried across from saccade to saccade.⁴ Grimes’ work prompted a renewed interest in the human ability to detect change and has inspired much speculation and experimentation on the change blindness phenomenon. It raises the question “Is the visual world a grand illusion?”⁵

In Dennett’s view, the whole phenomenology of perception is largely an illusion. The common conception is that there is a seamless stream of consciousness; Dennett’s view is

---

that consciousness is a fragmented, disjointed thing and that we piece it all together in retrospect. Vision researchers who set out to explain how we arrive at the seamless detailed, integrated, deep and colourful reality which we experience are actually falling at the first hurdle. Dennett maintains that the little we do see is largely in black and white, patchy and superficial. That is, we are not registering the world in anything like the detail that we think we are and we do not recall from moment to moment most of what presents to our retinas.

Dennett’s experiment with the playing card demonstrates this nicely. He comments:

The visual field seems to naive reflection to be uniformly detailed and focused from the center out to the boundaries, but a simple experiment shows that this is not so. Take a deck of playing cards and remove a card face down, so that you do not yet know which it is. Hold it out at the left or right periphery of your visual field and turn its face to you, being careful to keep looking straight ahead (pick a target spot and keep looking at it). You will find that you cannot tell even if it is red or black or a face card.\(^6\)

You will be aware of motion at the periphery of your visual field but will not be able to identify the card until it is almost right in front of your eyes. The reason that we cannot see colour or identify objects outside 5 degrees of dead centre is because only the receptors (cones) of the fovea around the centre of the retina generate a colour signal. The rest of the receptors (rods) cannot generate a colour signal. These facts are not remarkable – the remarkable thing is that we don’t notice this staggering deficiency in our vision. Given these, and other facts about the limitations of the human visual system (we have two upside down distorted retinal images and our eyes constantly dart around in saccade movements during which the retinal image “greys out”) it is also remarkable that vision research didn’t predict change blindness and inattention blindness earlier.

In fact, the results of the hundreds of change-blindness experiments that have been conducted since 1991 are so surprising that they have generated a new interest in the role of attention in both perception and consciousness. They have also given birth to a new discipline – visual metacognition – the study of people’s beliefs about vision. Research into visual metacognition is exploring how people not only fail to notice or attend to most of a scene, but always overestimate their ability to remember what they have seen. That is, we massively overestimate our powers of perception, attention and memory, and we similarly overestimate other people’s powers of visual cognition.

In his paper entitled “Is the Visual World a Grand Illusion?” Alva Noë summarises the significance of the change blindness work as follows:

The fact of change blindness is widely thought to have several important consequences. First, perception is, in an important sense, attention-dependent. You

---

\(^6\) Dennett, *Consciousness Explained*, 53-54.
only see that to which you attend. If something occurs outside the scope of attention, even if it's perfectly visible, you won't see it. In one study, perceivers are asked to watch a video tape of a basketball game and they are asked to count the number of times one team takes possession of the ball (Neisser, 1976; Simons & Chabris, 1999). During the film clip, which lasts a few minutes, a person in a gorilla suit strolls onto the centre of the court, turns and faces the audience and does a little jig. The gorilla then slowly walks off the court. The remarkable fact is that perceivers (including this author) do not notice the gorilla. This is an example of inattentional blindness.

Second, perception is gist-dependent. Some changes, for example, in the features that affect the gist of the scene, are more likely to be noticed (Simons & Levin, 1997). Third, it seems that the brain does not build up detailed internal models of the scene; that is, it doesn't perform the integration of information across successive fixations, contrary to the assumption of traditional orthodoxy (Blackmore et al., 1995; Rensink et al., 1997; O'Regan et al., 1999; Rensink et al., 2000; Noë et al., 2000). Or if it does, we have little easy access to this detail. If we did, then presumably we'd keep track of change better than we do.\(^7\)

The “basketball and gorilla” experiment to which Noë refers has now become part of the popular folklore of change blindness, and has been repeated on network television to the general dismay of unprimed viewers. Daniel Levin and Daniel Simons embarked on a series of increasingly surprising change blindness experiments in which subjects failed not only to notice changes of clothes and furniture in a scene but even failed to notice if one actor was switched for another. There was, at the time, some scepticism as to whether these change blindness results could be achieved using subjects in real world situations (as opposed to on video). In response to this scepticism, Levin and Simon developed a variety of situations where subjects were exposed to a quick substitution of their conversation partner (see fig.5). Levin comments:

We still found that about half failed to detect this change. In one case, an initial experimenter approached a subject on a college campus and asked for directions to a building on campus. Midconversation, two other experimenters carrying a door walked between the subject and the first experimenter. While the subject's view was briefly blocked, one of the experimenters carrying the door traded places with the first experimenter, who walked off behind the door. Thus the subject's conversation partner suddenly changed from one person into another. Yet, even though the change occurred right in front of them, about half of the subjects missed it, continued the conversation as if nothing had happened, and were later quite surprised to find that the person who finished the conversation with them was not the same person who had started it (Simons and Levin, 1998). In replicating this effect using a number of different scenarios, we have found that subjects also missed substitutions both when they were photographing the experimenters and when they were receiving consent forms from them (Levin et al., 2002).\(^8\)

---


The results of these trials amazed even the inventors of the experiment. It is hard to believe that one could fail to notice if one’s conversational partner was suddenly switched. It further confirms that we overestimate the level of visual detail which we notice and retain when we look at a scene. If indeed our visual system is working with such impoverished stimuli, are we justified in concluding that the visual world is a grand illusion? Noë says “no”. He cites the fact that despite the results of change blindness experiments we are still able to navigate the world successfully, and concludes that at the sensorimotor level at least, our visual system is giving us an adequate account of the world. I will return to the significance of the sensorimotor aspect of the visual system when I have analysed the significance of this change blindness work for a theory of depiction.

There are at least two major implications in the change blindness studies for a theory of depiction:

1. We seem to perceive a far less densely populated and detailed image than we imagine. It is therefore possible for an artist to trigger recognition with a picture which is far less detailed than a full-colour photograph or realist painting or drawing.
If we are radically mistaken about our visual phenomenology, intuition is not serving us well in our quest to understand how we see pictures, and the technique of introspection is also suspect. It is clear that only rigorous experimentation using the methodologies of psychophysics can illuminate how we see pictures.

It should, however, be remembered that we are not here attempting to solve all the mysteries of visual perception; we are merely trying to ascertain how an artist exploits the abilities (and disabilities) of the human visual system to evoke objects and scenes using pigment. Wooding’s studies of how artists use composition to direct our attention tell us something about how the configuration of a whole scene can lead the eye to significant areas. What we now need to discover is how our visual system recognises the subjects once the eye has been drawn to the relevant parts of the picture. That is, we saw in the Veronese that light, colour, gaze and presumably faces were the main attention grabbers. I’m interested in how our attention can be grabbed so quickly by key features of a scene such that recognition is triggered almost instantly. Podro, and to some extent Wooding, are interested in continuing recognition. I am primarily interested in how the interpretation gets a foothold. I suspect we need to understand the mechanism whereby we make that instantaneous rough judgement about what we are looking at, before we can analyse how recognition is sustained. I originally became interested in the change-blindness experiments because they seemed to confirm that we visually recognise things on the basis of very little or very poor visual stimulus. It seemed that this research might throw some light on how we can so easily recognise very sketchy line drawings or smudgy partial depictions such as Constable’s cows.

[In the next section of my thesis I develop a series of arguments and visual demonstrations of how work on object recognition, scene gist, change blindness, and the sensorimotor system has the potential to explain many of the puzzling features about how rough sketches can be so evocative.]

Bibliography
