

1 reading the visual

I am driving along in a car in the country. As I drive, I am looking out the windows—straight ahead, to the right and left, through the rear-view mirror—at the sky, hills, bush, road and the other vehicles around me. I am moving along this road, and through this landscape, at speed—say, 80 kilometres an hour. Everything I see is seen at speed: I am moving past trees, meadows, cattle and slower vehicles; and faster vehicles are moving past me. Even though I am travelling at 80 kilometres an hour with my vision framed, and thus partially restricted, by the mirrors and windows of the car, I can still see and negotiate my environment (road, trees, road signs, other cars). I drive on the road, in the slower lane, seven car-lengths from the vehicle in front of me. I observe speed signs, and change lanes when I come across a slower car without causing any accidents.

introduction

Suddenly a kangaroo jumps out of the bush, and bounds across the road—a not unusual occurrence around here. I'm alarmed—I know from experience what damage a car can do to a kangaroo, and vice versa. I rapidly focus my attention on the kangaroo, taking in its speed, size, trajectory, distance from my vehicle, and the rate at which I am approaching it. In almost the same instant I break and swerve to the side (somehow I know there are no cars around me), and miss it. I drive on, more alert, occasionally scanning the bush ahead for more kangaroos. When I arrive at my destination (my parents' house in the mountains, a place I have driven to many times), I have almost no recollection of the drive, apart from the incident with the kangaroo.

I have arrived at my destination, and I am taking a photograph of part of the house and the front part of the property (see Figure 1.1). I am taking the shot from the same level as the house but 20 metres to



Figure 1.1 A Lake, a Tree

the side, and I am only framing part of the house (which includes the verandah, a typical rural feature) so I can include two sets of trees. The first set is located just in front of the house, and the trees are leafless; the second set is another 60 metres down the slope on which the house sits, and the trees are solid with foliage. The backdrop to the house and trees is a thick mist

which has partly covered the lower trees, and seems to be moving towards the house.

I have a few things in mind which have led to this arrangement of the shot. I want to produce a sense of space (the house as one small part of a much larger property, which is one of the reasons I have included the second set of trees). I want to catch the property as it usually looks at this time of year. But I also want to emulate those landscape paintings and photographs which contextualise signs of human presence (the house) within the forces, power and rigours of nature (the trees and the enveloping mist). The leafless trees are situated at the centre, and take up almost half of the photograph, while the house is peripheral (and consequently relatively insignificant). My focus will be on the objects in the foreground (particularly the tree branches, and the way they tower over the house); the rest of the scene (the solid trees, the paddock, the mist) will be slightly blurred.

the activities of seeing

Planning and taking a photograph is, like many human activities, an intensely visual experience; so is driving a car, where we are constantly visualising and making sense of the space through which we are moving. There is one big difference, of course: driving a car is a relatively unreflective activity and even below the level of consciousness, while taking a photograph is usually conscious, deliberate and self-reflective. In other words, we usually pay a great deal of attention to what we are doing when we are photographing a scene; but when we are driving a car we are often doing so on automatic pilot, and only pay close attention to what is around us when we need to (for instance, when a kangaroo jumps out of the bush or when we are looking for a place to stop and have lunch).

This difference between the two activities—a difference of levels of attentiveness, among other things—is one of degree rather than of kind because, whether we are aware of it or not, in both instances we are making (that is, actively ‘bringing about’) the (visual) world around us. When driving a car, or arranging a photograph, we are not simply seeing and taking in everything that is available to our range of vision. The space I photographed contained an extraordinary, almost infinite, amount of detail that I simply didn’t see. There may have been rabbits, camouflaged and keeping still on the slope; birds blending into the branches of the trees; various plants and types of grass around the house; kangaroo dung by the lower trees; the front roof of the neighbour’s house poking through between the bare trees and my parents’ house; a small puddle created by a dripping tap, and many other details. Some of these things are more or less visible in the photograph, but the rest weren’t seen and haven’t been shown. Had I seen them, I might have changed the angle, distance, speed, frame and focus of my shot, and produced a different photograph (‘puddle outside the house’, ‘rabbits in the paddock’). But every act of looking and seeing is also an act of not seeing, even when we are being attentive.

This is true to an even greater extent with the act of driving a vehicle. It seems strange to suggest that I can be more attentive and reflective when taking a photograph, which is a relatively trivial activity with no serious consequences (about the worst result would be that my parents dislike the way the house is shown, or maybe I could get the focus wrong) than when driving a car, where one wrong move could cost me my life. But a lot of our visual activity in driving is more or less automatic: we see where we’re going and what is around us, but our attention is usually focused on only one or two spaces (the lane we’re driving in, the car in front of us). And even here our attention is often more general than specific. We make sure we’re driving within the lines that designate our lane, but we don’t usually look to see whether the lines are all the same length, or partly worn away; or notice the texture, condition or colour of the surface of the road (oil stains, small cracks, small tufts of grass, squashed cigarette packets). And the car in front of us is often seen in a very indistinct way. We might be aware of the distance between the two vehicles, or the speed, size and colour of the other car, but we rarely look at it in a detailed way, and might be hard pressed to recall its make, year of production, condition of the tyres, number of people riding in it, or their gender, age and skin colour.

There are other reasons why we might not pay as much attention when driving as when taking a photograph. The trip might be over several hundred kilometres, and take hours. We simply can't look at things in a detailed and attentive way for that length of time, particularly when we're moving at speed. And moreover, while there is a link in photography between attention and enjoyment (we have chosen to look at things, frame them and capture them on film), a car trip is more often a means to an end rather than an end in itself (I drive to get to work, to visit my parents, to go to a shopping plaza). In other words, it is in my interest to be attentive when driving only insofar as my or somebody else's safety is concerned (watch to ensure that I'm not exceeding the speed limit, and that I'm travelling within my lane), for reasons of economy (I only have a limited amount of attention to give), practicality (I'm moving too quickly to take most things in), and in order to ensure that I achieve what I set out to do (get somewhere where I can see my parents, or take photographs, or go shopping).

seeing as reading

We have covered three main points so far. Firstly, when we see things we are actively engaging with our environment rather than simply reproducing everything within our line of sight. Secondly, every act of looking and seeing is also an act of not seeing—some things must remain invisible if we are to pay attention to other things in view. Thirdly, the extent to which we see, focus on and pay attention to the world around us (the three actions are inextricably linked) depends upon the specific context in which we find ourselves.

While the process of making and negotiating the visual (whether driving a car or taking a photograph) is always informed by the notions of attentiveness, selection and omission, and context, there are other issues which we need to consider, such as when we do focus on, attend to and see something, and why do we see things differently over time, or from other people?

Consider the first paragraph in Stephen Crane's short story 'The Open Boat', which is about the experience of four men who take refuge in a rowboat after their steamer has sunk:

None of them knew the color of the sky. Their eyes glanced level, and were fastened upon the waves that swept towards them. The waves were of the hue of slate, save for the tops, which were of foaming white, and all of the men knew the color of the sea. The horizon narrowed and widened, and dipped and rose, and at all

times its edge was jagged with waves that seemed thrust up in points like rocks. (Crane 1960: 140)

The men in the boat don't know or see the sky because their attention is focused on something of more immediate interest: the waves that threaten to overturn or smash their boat, and take their lives. They see the waves in great detail: they are 'the hue of slate', with foaming tops, and they seem sharp and threatening like rocks—that is, the waves are the same colour (and, by extension, hardness) of slate, and as the boat comes down upon the waves it appears to be landing on sharp, hard rocks.

Now we could say that the psychological state of the men in this extreme condition has produced an effect so that the waves have become, in their minds, like rocks. But, as we saw with our previous examples of driving and photography, every act of perception takes place within a context that orients, influences or transforms what we see. Observing a kangaroo from the balcony of a café at a nature park produces a very different sight from what we experienced when we swerved to avoid one on the road. Watching the approaching mists when we are deep in the bush with a broken ankle and unsure of our way home is a very different experience from that of treating the mist artistically, as an aspect of a photograph that depicts natural forces. And when watching a storm from the safety of a cliff we may see the slate-hued waves, and thrill to the drama and tension of the scene, but this does not equate with how the sea appears to Crane's men in that open boat.

Every perception and meaning is the product of psychological, physiological and, above all, cultural contexts (I'm stressed, I'm not wearing my glasses, I'm lost, I'm an artist). In other words, the things we see aren't simply 'out there' in any ideal or unmediated way; rather, we understand, evaluate and categorise—that is to say, see—things in terms of a set of resources that we take from our cultural contexts. It has long been accepted in what we call the human sciences and the humanities—particularly in disciplines such as sociology, anthropology, linguistics, philosophy, literature, psychoanalysis and cultural theory—that we make sense of our world through the different meanings, ideas and categories available to us. And it is this situation of a culture more or less seeing through and for us, combined with the inflection or influence of different psychological and physiological states, and of-the-moment contexts, that produces what we see.

We can carry this insight further by suggesting that when we see we are, in effect, engaged in an act of reading (the visual). When we read a book we try to follow, consider and understand the material at hand (the words, the sentences, the story), and we end up making both meanings and connections between different meanings. In Jules Verne's *20,000 Leagues Under the Sea*, for instance, we come to understand that Captain Nemo is keeping Professor Aronnax and his companions prisoners aboard the *Nautilus*, and that he is obsessive about not returning to land; we infer that Nemo has suffered some great psychological hurt or loss, and that he will never let them leave the giant submarine alive. We could say that the story of the book is about the relationship between two different sets of wills, and how this is played out (will the *Nautilus* destroy other ships? Will Aronnax and his friends thwart Nemo or escape?). But no two people will read the book in exactly the same way: some readers will see Nemo as a heartless murderer, while others will see him as rightfully enacting revenge on a world that robbed him of his wife and children, who were killed in a naval battle. The point is that the same book will be subject to different readings and interpretations precisely because people approach it from different backgrounds and perspectives.

There is another reason why the book will be subject to different readings: readers will want different things from it. A person with two hours to devote to a rollicking adventure will read it differently from someone studying the book for a school or university exam. Roland Barthes writes in *The Pleasure of the Text* that, when he has a story in front of him, 'I read on, I skip, I look up, I dip in again' (Barthes 1975: 12). And he refers to 'two systems of reading: one goes straight to the articulation of the anecdote, it considers the extent of the text, ignores the play of language'—if I read Jules Verne, I go fast—while the other reading 'weighs, it sticks to the text . . . [and] grasps at every point' (Barthes 1975: 12).

These descriptions of different ways of reading a book could just as easily be applied to practices and ways of seeing. Barthes' reference to his 'skipping and dipping' style of reading, for instance, pretty much sums up the orientation of the car driver who takes in the (visual) bare minimum, while the reader who 'weighs things' and closely examines the text is like the photographer carefully attending to and considering everything within the photographic frame.

When we read a book there is always a context to that act of reading; we might, for instance, try a book because we are familiar with the author's other works or critical reputation, or simply because we wanted to pass the time with a 'quick read'. But even if we had never

heard of the book or the author, we have access to other signs, such as the title, which would help us categorise—and thus prepare for—what we were about to read. We would probably expect *20,000 Leagues Under the Sea* to be an adventure story rather than a scientific study of deep-sea life simply because we know that adventure stories have titles that refer to exotic, dangerous and far-away activities and places, while scientific works are much more specific about their subject, and the language used is usually less accessible (for instance, 'Protandry and the Evolution of Environmentally-Mediated Sex Change: A Study of the Mollusc' is clearly not an adventure story). Similarly, everything we look at and make sense of, whether it is a photograph or a set of objects within our purview, comes with a history of commentaries, meanings and annotations which disposes us to read it in a particular way.

The relationship between those forces which dispose us to categorise and see the world in certain ways, and the kinds of visual texts that subjects make, can be usefully explained through reference to two contexts—one taken directly, and the other extrapolated, from the work of Pierre Bourdieu. The first is the habitus, and the second is cultural literacy. Bourdieu famously defines the habitus as 'the durably installed generative principle of regulated improvisations . . . [which produce] practices' (Bourdieu 1991: 78). In other words:

Habitus can be understood as a set of values and dispositions gained from our cultural history that stay with us across contexts (they are durable and transposable). These values and dispositions allow us to respond to cultural rules and contexts in a variety of ways (they allow for improvisations), but these responses are always determined—regulated—by where we have been in a culture. (Webb et al. 2002: 36–7)

Our cultural history and trajectories naturalise certain values and ideas, and effectively determine our worldview—that is, they predispose us to see and evaluate the world in certain ways. Central to this is what Bourdieu terms distinction: this is tied up with the notion of taste, which generally means having a refined, educated, sophisticated and aesthetic worldview, rather than simply seeing, evaluating and categorising things 'naively' (say, in terms of their use value). A good example of distinction as it manifests itself in everyday life is this story about the philosopher Ludwig Wittgenstein, who was taking a walk in his garden. His gardener was surprised to see him, and said

'Professor, what are you doing here?' to which Wittgenstein is supposed to have replied, 'What are any of us doing here?'

In order to see how distinction, and habitus, influence the way people see, let's consider Figure 1.2, a photograph of a pile of dishes and utensils, presumably left sitting in the kitchen. According to Bourdieu, a sophisticated habitus would be perfectly capable of seeing—in fact, might be disposed to see—what was within the frame as something other than a simple reproduction of a domestic scene. A so-called sophisticated eye might see carefully arranged patterns and motifs (note how the utensils are all pointing one way), references to other texts (the wobbly pile of dishes as a Tower of Babel) and social commentary (the dishes as an index of the chaotic state of modern life). People who did not share this habitus might dismiss such readings as boring or pretentious. They could respond that everybody knows what the kitchen is like in a shared house, and what an unwashed pile of dishes looks like—and they are neither beautiful, nor capable of saying anything meaningful.

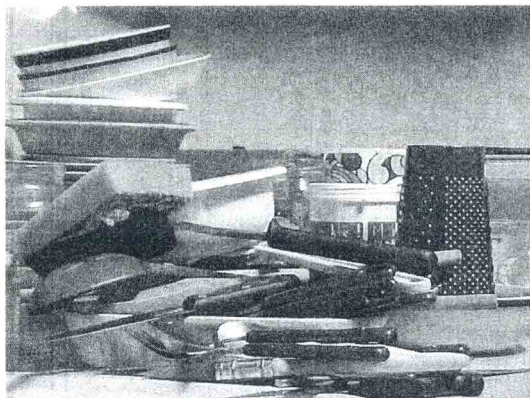


Figure 1.2 Everything but the . . .

Distinction would not only dispose people to see this photograph as meaningful or beautiful, however; it would also supply them with the knowledge and ability that would make such a perspective possible, and provide it with legitimacy. The name we give to this combination of knowledge and skill is *cultural literacy*. When we think of the notion of literacy, we usually associate it with reading and writing skills but in this case the term refers to a general

familiarity with, and an ability to use, the official and unofficial rules, values, genres, knowledge and discourses that characterise *cultural fields*. *Cultural literacy in this sense is not just familiarity with a body of knowledge; it also presupposes an understanding of how to think and see in a manner that is appropriate to the imperatives and contexts of the moment.*

One way to demonstrate what we mean by this is to give an example of a lack of literacy. In the Marx Brothers' film *Animal Crackers*, Groucho Marx is playing a famous explorer, the guest of a

woman (played by Margaret Dumont) who is a member of high society. Another guest is a rich art collector who is showing his most recent acquisition, the famous painting 'After the Hunt' by the French artist Beaugard (a painting which shows a hunter and his hounds). The painting has been stolen and replaced by a copy. When it is unveiled, the owner laments that this is not the original. When asked how he can be sure that it is a fake, his reply is that anyone can see that it is a poor imitation. He is about to explain what he means, presumably by referring to the poorly drawn details and inferior brushwork, but before he can do so Groucho adjusts his glasses, focuses on the painting and says, 'My God he's right. One of the dogs is missing.' The owner was showing his art literacy by being able to distinguish, at the level of technique and detail, between a masterpiece and a forgery. Groucho shows his lack of literacy by introducing a naïve, content-related issue ('There's one less dog!'). Of course, Groucho's illiteracy has a point to it: it demonstrates that while a 'high art' habitus allows people to see particularities like texture or subtleties of light, it also blinds them to what is in front of their noses. The owner presumes that a forger will at least get the details right (only one man on a horse, precisely five hounds). But Groucho, who has a more practical way of looking at things, sees what is obvious to him but hidden from the supposedly sophisticated eye.

This very peculiar and seemingly naïve (to the owner and the rest of the party) way of looking at and making sense of art is not in any sense idiosyncratic or accidental. Put simply, the cultural contexts, fields and institutions that Groucho inhabits and moves through (what we call his 'cultural trajectory') now effectively 'inhabit' him, influencing and determining what and how he sees. If Groucho were to become part of this more 'sophisticated' circle, their way of seeing would become more natural to him, and eventually influence, even determine, what he sees.

Our situation is pretty much the same as Groucho's, **seeing in conte** in that what we see is inextricably linked to, and is a product of, our cultural trajectories, literacies and contexts. This applies even when we see something for the first time. Given that we know, think and see within our cultural frames, a truly 'new visual experience' is almost impossible to imagine. Even if we were subjected to something literally 'out of this world', like being abducted by aliens and taken to another planet, we would still see by using the categories and forms of evaluation that characterise

our habitus. This would happen partly because we set up and make use of distinctions such as human/alien, even though we've never seen a 'real' alien. Of course, we have seen representations of aliens, which are normally distinguished from humans by their colour or size ('little green men'), body parts (Zaphod Beeblebrox, in *The Hitchhiker's Guide to the Galaxy*, had two heads which were always arguing) or supernatural abilities (Superman can leap tall buildings in a single bound).

The process through which we might see an alien—real or imagined—is more or less the same as what is involved when we categorise 'real people' or groups within our culture. We start off with a binary concept (human/alien) which is 'filled out' by various signs (humans walk upright and have one head), and when we come across something or someone which doesn't fit into one part of the binary because of an excess (an extra head) or lack (no head at all), we simply categorise, evaluate and see them (remember, this all happens more or less simultaneously) as the other part of the binary (the alien).

There is another interesting aspect to this question of what happens when we see something 'for the first time', and the answer is that often... we don't see it the first time we look at it. The musician Tom Verlaine, formerly of the New York 1980s punk band Television, sang in his song 'Postcard from Waterloo' that 'I recall the actor's advice/That nothing happens til it happens twice'. The point he is making is that the first time something appears which doesn't obviously correspond to categories with which we are familiar, or which we don't expect to see, we are likely to miss it. Using a similar logic, science historian Thomas Kuhn writes that the physicist Roentgen's 'discovery' of X-rays was at first:

greeted not only with surprise but with shock. Lord Kelvin at first pronounced them an elaborate hoax. Others, though they could not doubt the evidence, were clearly staggered by it. Though X-rays were not prohibited by established theory, they violated deeply entrenched expectations... By the 1890s cathode ray equipment was widely deployed in numerous European laboratories. If Roentgen's apparatus had produced X-rays, then a number of other experimentalists must for some time have been producing those rays without knowing it. (Kuhn 1970: 59)

How do we come to see what we have been overlooking? Kuhn writes that, in the scientific field: 'Discovery commences with the awareness

of anomaly... with the recognition that nature has somehow violated the paradigm-induced expectations that govern normal science' (1970: 52–3). In other words, our habitus disposes us to see certain things, but occasionally there is a misfit—or an anomaly—regarding what we expect to see and what we visually 'register'. Once this anomaly is repeated, we might start to reconsider what it is we are seeing—or overlooking.

We can exemplify this by returning to Verlaine's reference to the 'actor's advice' about things needing to happen twice. What this means is that we sometimes fail to see the significance of something until we are aware of what we could call a pattern. So, in Peter Jackson's film *The Fellowship of the Ring*, the hobbit Bilbo Baggins is represented as an inoffensive, generous and altogether nice type who seems untouched by desire, passion or greed. But he has a secret: he owns a ring that has cast an evil spell on him. We see signs of this when the wizard Gandalf asks him to hand over the ring, although the first few manifestations (a slight hesitation in responding to Gandalf's request, a strange look on his face as he ponders what to do next) could easily be overlooked. It is only when his determination to keep the ring leads him to act 'out of character' (he becomes suddenly violent and irrational), and when his face is completely transformed by the power of the ring (his features become contorted with rage), that we notice the pattern and understand the secret—he is possessed by the power of the ring. If we are familiar with Tolkien's story before viewing the film we will expect this to happen, and see what is happening 'the first time'; if we aren't, however, then Bilbo's hesitation and odd looks will just be part of a Kuhnian anomaly—until we perceive them as part of a pattern.

Up to this point we have concentrated on explaining how and why

people see in particular ways, and we have referred to habitus, cultural trajectory and cultural literacy as the most important factors in determining what we see. But we also suggested that, whether our seeing is conscious or unconscious, the process of reading the visual relies on the same techniques. The techniques we will consider include selection, omission and frame; signification and evaluation; arrangement; differentiation and connection; focus; and context. It is important to keep in mind that there is no necessary temporal distinction between these techniques: they are part of the same process of making the visual, and one cannot be conceived without regard to the others.

techniques of seeing as reading

general principles process

The first and most important techniques of reading the visual are selection and omission; as we pointed out earlier, every act of looking and seeing is also an act of *not* seeing. Consider the text shown in Figure 1.3, a photograph of a woman sitting on the steps of a house with a dog. The photographer probably had a considerable amount



Figure 1.3 Lassie, c. 1920

of material to play with; there is a house, perhaps extensive gardens, a lawn, other people and animals, a street, other houses and maybe some cars. The photograph only shows us a selection of these: it includes a woman, a dog, the steps, some flowers or bushes, the lower part of the door and a shuttered window. We could consider a number of other aspects of the selection/omission process, such as the fact that we can see the woman's boots, but not her eyes.

The selection of these details (and the omission of the others) helps to constitute and make the visual. It doesn't matter whether we are looking at a photograph, painting or a street scene: by paying attention to and focusing on the woman and her dog in this space, the viewer effectively constructs a frame around the scene. This is productive in two ways. Firstly, it suggests a set of relationships between, and stories about, the various parts—perhaps the woman is playing with her dog at her house; perhaps she is simply relaxing on her steps. Secondly, it establishes a (usually temporary) hierarchy with regard to the potentially visible: that is to say, whoever took this photograph or observed this scene decided (at a conscious or unconscious level) that this content within this space at this time was interesting or worthy of attention. In other words, they made an evaluative decision. This may have been careful and deliberate (they

many act of photography is an act of A. red

of a house, perhaps extensive gardens, a lawn, other people and animals, a street, other houses and maybe some cars. The photograph only shows us a selection of these: it includes a woman, a dog, the steps, some flowers or bushes, the lower part of the door and a shuttered window. We could consider a number of other aspects of the selection/omission process, such as the fact that we can see the woman's boots, but not her eyes.

The selection of these details (and the omission of the others) helps to constitute and make the visual. It doesn't matter whether we are looking at a photograph, painting or a street scene: by paying attention to and focusing on the woman and her dog in this space, the viewer effectively constructs a frame around the

set the scene, posed the woman, took the shot) or spontaneous (they were wandering by, the scene appealed to them, they took a photograph). Either way, these acts of selection, omission, framing and evaluation produce a visual text.

What do we mean by the term 'text'? Usually we think of a text as something like a book—that is, it is an object that consists of words on pages, sometimes accompanied by photographs, divided into chapters, authored by someone, with a title and a cover. Sometimes we extend this to cover other mediums, such as film and television. The defining characteristic of a text is that it is (or is treated as if it were) a unit: we think of a book or a film as a text because the various parts are both related and bound together. This occurs through an arrangement of what we call signs, which can be defined, very generally, as something (say, a word or a photo) that is read as meaningful by someone (a reader or photographer). A group of signs is being read or treated as a text when someone considers all the signs as a unit. For example, we usually treat a book as a text which is made up of signs that possibly include the name of the author, the colour of the cover, the title and the publisher's name. In other words, a text is made up of signs that are considered to exist in relation to other signs, the sum of which is denoted by a frame of some kind.

f. one
D. one
Shr.

The most important point to keep in mind about this definition is that texts are not simply objects which always retain the status of 'text'. Rather, texts are produced or created; this process of production is an ongoing one; and the status of signs and texts is always relational and contingent. In other words, there are no natural units of signs within cultures—or anywhere in nature, for that matter. Every time we treat something as if it were a text, we create a unit out of an infinite number of potential signs. So the person who took the photograph of the woman sitting on the steps took a number of potential signs (the woman, the dog, the steps, the house, the door, the garden) and included them within the frame of the camera's lens, and subsequently the frame of the finished photograph. Exactly the same kind of process occurs when someone walks up to a scene and notices certain details within the frame of focused, attentive vision.

There is (at least) one major difference between these two visual activities. When we look at a photograph, television, film screen or painting, we normally apprehend something that is in front, distanced and detached from us, whereas the texts we create as we see the world are all around us, like the visual equivalent of surround sound, or virtual reality; we are located within them and they in us. So there

are several technical tasks we have to perform, consciously or otherwise, in order to stitch together (what Jacques Derrida in *The Truth in Painting* calls 'suturing') all these elements so that they appear to be a single, continuous visual world.

Two important factors here are *attention* and *focus*. If we are attending closely or carefully to an event, person, thing or scene, we will create a text that is made up of what we can call *contiguous elements*. So if we were staring out of a window we might see tree branches waving in the wind directly in front of us and a cloudy sky above, but we would also be likely to include the window and curtains or blinds, the computer that is partly between us and the window, a section of the desk on which the computer is sitting, the telephone and the pile of books slightly to the side. We might be more peripherally aware of other objects within our purview, such as the walls of the room, bookshelves, papers, carpet or the ceiling. Our eyes may be caught by the colour or movement of things—the deep purple of the walls, the brightly coloured, whirling images of the screen saver on the computer. But the decision about what is included within the main frame and what is left to the periphery is very much of the moment. In other words, if I watch the computer screen or look out the window, the function or context of my looking and seeing (whether to do something specific like check email, or just to look dreamily away from my work) will determine what is included in the visual texts I produce.

seeing in time and motion

A number of elements contribute to or facilitate the process of suturing the world to make a text. *Colours* help us to differentiate elements within our purview (the green of the trees and the blue of the sky); so do *shape* and *movement* (the still, rectangular window, as opposed to the relatively amorphous, waving tree branches). The use of colour, shape, movement and other elements (such as *texture*, *distance* and *light*) does not occur in an unmediated way, however; rather, the *extent* to which, and the *how*, we recognise, know and use them are tied up with our immersion in, and relationship with, our *cultural world* and its categories. We need to bear in mind that notions such as colour, shape and texture are culturally specific; we naturalise the world, give it stability and coherence, and are able to understand and explain it through our ability to maintain the (optical) illusion that what and how we see, and the texts we create, are real.

Of course these visual texts are ephemeral. In a sense they are never stable or really 'themselves'; after all, as I watch the trees, time

is passing. I am changing and so are the trees. This is even more pronounced, and the texts that are produced are far more impressionistic, when we or the world are moving at speed. Even if I am keeping still, the slightest movement of my eyes or head automatically changes what is available to me to be processed, combined, textualised and read. If I look slightly to the left, I might notice a bird on the window ledge, a poster on the wall, a brightly coloured paperweight—and so I make a new text. Moreover, the dreamy state that induced me to look out the window might be replaced by a more focused mood, set off by some detail: I notice the paperweight keeping down a pile of papers I have to read, so I suddenly focus on them.

We suggested that this process of production is an ongoing and transformational one, and the status of signs and texts is always relational. Let's consider the first of these points, using the example of the photograph of the woman and dog to which we referred earlier. The photographer who took that shot could have taken a second photograph from 10 metres further away, this time including the whole of the house, more of the garden, and fourteen other dogs standing in the doorway. This would have produced a different text because the potential relationship between the various signs (and therefore the meanings available to anyone making sense of the text) would have been considerably altered; in the first shot, for instance, the woman dominates space, but she would be only one small part of the second shot. But the same principle (that *remaking a text always transforms the text*) would apply even if the photographer stood on the same spot and took another shot from the same angle, with the same frame. Why? Because time will have intervened in some way. In the most obvious case, some additional detail will have moved into the space of the frame—for instance, the fourteen dogs originally standing just outside the frame could all have run down and jumped on top of the woman. This would produce a different set of relationships between the signs, a different set of meanings and a very different text (say, to being potentially comic). But even when there are no new signs, the original signs within the frame will have changed in some way (the woman may have noticed the photographer and smiled, for instance).

Gilles Deleuze draws attention to this issue in his discussion of Henri Bergson's theorising of the relation between movement and instant (that is, time). Bergson puts forward the proposition (a paradox of sorts that he takes from the Greek philosopher Zeno) that movement and instant are both inextricably linked to, and inexplicable without regard to, one another. At the same time, neither is real

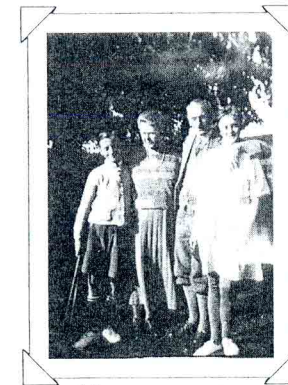
in any sense—they are artifices or illusions of perception. As Deleuze writes:

You cannot reconstitute movement with positions in space or instants in time . . . You can only achieve this reconstitution by adding to the positions, or to the instants, the abstract idea of a succession . . . And thus you miss the movement in two ways. On the one hand, you can bring two instants or two positions together to infinity; but movement will always occur in the interval between the two, in other words behind your back. On the other hand, however much you divide and subdivide time, movement will always occur in concrete durations . . . thus each movement will have its own qualitative duration. (Deleuze 1986: 1)

In other words, no matter how quickly we look at the same scene of the woman and the dog again and again, and no matter how much we are given to believe that things are the same (because it seems to be the same text), things have changed in the intervals of perception. But we can't capture the movement of time (the changes that differentiate one text from its successors, such as the woman beginning to smile, or the dog looking up) because the text is only available to us as a frozen instant, a text in time.

The second point we made is that the relationship between, and the status of, signs and texts are always relational and contingent. What do we mean by this? We suggested that a text is made up of different signs considered and framed as a unit—a woman, a dog, some steps together forming a photograph. But what if the photographer, or someone watching the scene, ignored everything except for the woman, so that she alone was in the frame? We could now say that, whereas before the woman was a sign in a text, now the woman had become the text, and the various details regarding her body and clothes (her boots, trousers, jumper, face, hair, hands) constituted the signs that made up the text. This process could continue almost ad infinitum. If we focused on the woman's face, it would become the text, and her eyes, nose and mouth would be the signs.

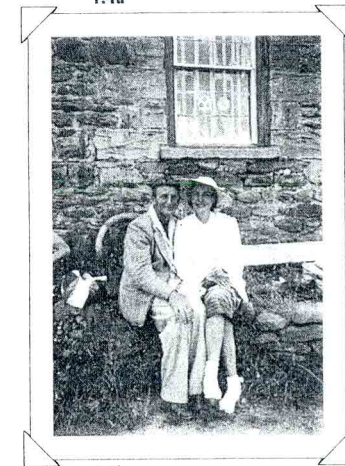
Figure 1.4 provides a good example of the relational character of signs and texts. There are five photographs arranged on a single page. They were not originally taken as a series, intended to be placed together, or considered for public consumption; rather, they were private family photographs which we have put together, not entirely arbitrarily, to make a text.



1.4c



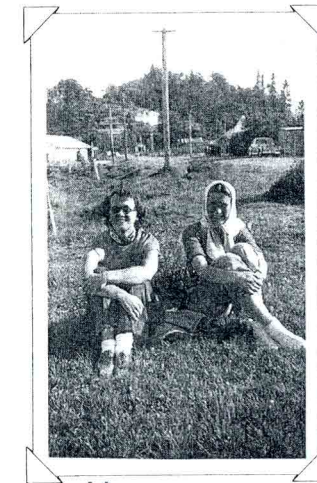
1.4a



1.4b



1.4d



1.4e

Figure 1.4 Five found photographs, Wellington, New Zealand, 2000

We can say that the combining of these particular photographs was not arbitrary because they are clearly connected through content (members of a family appear and reappear in them) and social function (they are all identifiable as being family portraits or photographs).

Each of the photographs can be read as a sign within the larger text. We could focus on one of the people (the woman in Figures 1.4a, 1.4b, 1.4c, 1.4e; the man in Figures 1.4b, 1.4d) and read between and across these signs, relate the content of each to that of the others, and produce a narrative or account of their lives (from youth to middle age, say). We might take this text as being about the family and its history, which would involve identifying the different generations and their relationship to one another, through reference to features such as clothes and physical characteristics. Each of these photographs-as-signs would have meaning in relation to the way they were read and contextualised with regard to the other signs—that is, their meanings and status would be determined by their textual place. But there is nothing to stop us from considering each of these photographs individually as collections of signs—that is, as texts. So the largest of the photographs (figure 1.4e) has a plethora of potential signs (the two women, their facial expressions and poses, the space between them, the rural setting) which can be collected together and read in relation to one another as a text without referring to any of the other photographs.

A sign, we have suggested, is anything that is treated as a meaningful part of the unit that is the text. We identify signs and group them together as if they were a unit by a process of relating available material to the other texts and text-types with which we are familiar from our memories and cultural histories. The use of other texts to create new texts is called *intertextuality* and the term for text-types is *genre*. In order to consider these two concepts and how they inform or influence visual activity, let's look again at the series of photographs in Figure 1.4.

We made the point that every photograph in the collection is made up of potential signs (the people, their clothes, their facial expressions and poses, the space between them, the setting) that could be treated as individual texts—without needing to refer to any of the other photographs. But when we consider that text with regard to one or all of the other photographs, we are making use of intertextuality—which means the process of making sense of texts by reference to other texts, or to meanings that have already been made in other texts. Let's look at Figure 1.4e. We can identify two obvious

signs—the two women in the foreground sitting side by side, smiling. We don't know anything about their relationship or their histories, so how do we make sense of or read those two signs? We can do so intertextually—by looking at some of the other photographs. The woman on the right of Figure 1.4e appears also in Figures 1.4b, 1.4c and 1.4d—or at least we presume that is the case because of the physical similarities between the various women in the shots. In Figure 1.4b, the woman is sitting very close to a man about her own age who probably has his arm around her waist or back. In Figure 1.4c she is much younger—perhaps in her mid-teens—and is posing with two adults and a boy, perhaps her parents and brother. In Figure 1.4d she looks slightly older than the woman in Figure 1.4e; she is at some kind of social function, and is the only person in the shot.

Taking these photos as a collection, we can 'read' a kind of narrative of the woman in Figure 1.4e. We have the woman-as-girl, growing up in what looks to be a middle- or upper-middle-class family in the first part of the twentieth century (which we identify from the hair-styles and clothing). She had a relationship with, and perhaps was married to, the man in Figure 1.4b. She was fair-skinned and probably lived in a sunny country (in three of the four shots she is wearing head covering of some kind, although this may be explicable in terms of the clothing fashions and conventions of the time). She probably grew up and lived in the country, rather than the suburbs or city. The young man we take to be her brother in Figure 1.4b is carrying rifles, and the houses and physical environments in Figures 1.4b, 1.4d and 1.4e all have rural characteristics (the rough stone material of the house in Figure 1.4b; the water tank in Figure 1.4d; the forest and sparsely housed scene in Figure 1.4e).

We might search these photographs for signs that would enable us to generalise about her history and life (she grew up and lived her life in the country in the mid-twentieth century), her nationality (Australian, New Zealander or South African), her predilections (in all the photographs she is wearing white shoes) and many other things. Some of these generalisations might be relatively obvious (for instance, the relationship between her clothes, her age and the approximate period in which she lived), while others are little more than conjecture (were the couple in Figure 1.4b married? Does the presence of head-wear mean she was sensitive to the effects of sunlight?).

We could go on like this indefinitely, bringing in new intertexts that change the way we read the photograph we originally considered (Figure 1.4e). The important points, however, are that we can and do read texts such as Figure 1.4e intertextually and, even when

we don't know specific details about those intertexts, we are disposed and able to make sense of and read them (we presume, without knowing for certain, that Figure 1.4c is a family shot showing parents and children). We are able to do this because every reading of a text is informed and influenced by our intertextual reference to and knowledge of the text-types that characterise our culture—what we call genres.

texts and genres

Genres—which we discuss in more detail in Chapter 4—can be defined as 'text types which structure meanings in certain ways, through their association with a particular social purpose and social context' (Schirato and Yell 2000: 189). We normally think of genres in terms of cultural fields and mediums such as fiction or film—for instance, detective, science fiction or romance novels; and action, horror or erotic films. Each of these genres is identifiable in terms of its content, narrative, characterisation, discourses, values and worldviews. A detective film will usually have a certain kind of content as a constant (a crime, or an act of violence), which will require the intervention of a detective who will investigate the scene, question suspects and take testimonies from witnesses; hunt for, find and analyse clues; and eventually uncover secrets, overcome the criminals and solve the crime.

There will, of course, be variations across these texts. A film of one of Conan Doyle's detective stories will represent Sherlock Holmes as detached, observant, attentive, analytical, incorruptible, well-mannered and supremely self-confident—all of which is shown in the way Holmes moves, speaks, looks and acts. The values and worldviews represented in the film (if they are faithful to Conan Doyle's original fictions) will usually be socially conservative, and pretty much in keeping with the dominant social values of the time and place (so servants will be treated as if they are naturally less valuable, interesting and intelligent than members of the middle or upper classes). Finally, the descriptions of places and events will be strongly informed by what we could call scientific orientation: rooms, furniture, spaces and people will be described and shown in careful, precise detail.

Not all detective films—or plays, cartoons or video games, for that matter—reproduce or partake of all of these conventions. The so-called 'hard-boiled' detective films made from novels written by American writers such as Dashiell Hammett and Raymond Chandler will vary or even repudiate some of these characteristics (most obvi-

ously, the detective might be a drinker and gambler who may become sexually involved with suspects, and may work outside the law). But, by and large, there are enough constants and carry-overs from Holmes to the hard-boiled detective stories (and later on, to the forensic, feminist and historical detective forms) for us to categorise them as belonging to the same genre.

Who decides to which genre a book or film—or any other text—belongs? Michel Foucault's work alerts us to the ways in which ideas, worldviews and categories (of people, thought or texts) are institutionally produced. That is to say, different fields and the institutions within them produce (authorised) knowledge and statements through which we see, categorise and make sense of the world. With regard to films, for instance, a variety of fields (academe, the media, government) and experts (film studies academics, film critics, politicians) will analyse, and make pronouncements about, a film's value and genre; and this will effectively determine where a film is shown (mainstream, art house or 'adult' cinema), what rating it receives, and therefore who is allowed to see it. Moreover, these comments, classifications and ratings will also orient the way audiences understand and evaluate a film.

The films around which these questions and issues are often played are those which have explicit sexual content—which usually means they will be categorised as pornography. But sexually explicit films are sometimes given a rating which allows them to escape this classification. For example, two French films with explicit sexual content were shown in Australia over a period of three years—*Romance* (directed by Catherine Breillat, released in 1999) and *Baise Moi* (Coralie Trin Thi, 2002). *Romance* was not finally classified as pornography, despite the fact that there were scenes showing actual sexual activity and implied depictions of sexual violence. The main reason the film received a restricted classification was because it was considered an art film: the director was known to be interested in, and had dealt with, philosophical, political and social issues (the nature of desire, masculine violence, the dehumanised state of modern society), and these same 'serious' issues were talked about and represented in the film. If we go back to our definition of genre, we see that it refers the 'particular social purpose and context' of text types. *Romance*, presumably because of the status of its director, was considered to be showing sexual content in order to explore contemporary social issues. In other words, unlike pornography, the sex scenes (as far as the national censorship board was concerned) had an artistic, social and educational

function, rather than being intended simply to produce sexual arousal.

Baise Moi was released in Australia three years later. Again, the film seemed to be oriented towards—or at least to be informed by—social criticism and artistic features, and it was being shown as an art house film. But after originally giving the film a restricted classification, the censorship board reversed its decision, which meant that the film had to be withdrawn from cinemas. On the face of it, there was very little to differentiate the film from *Romance*, but in three years the social and political climate had changed sufficiently for the two films to be given entirely different classifications and categorised as separate genres—*Romance* was effectively categorised as an art film, *Baise Moi* as pornography.

These kinds of official classifications—and effectively generic categorisations—of texts influence and orient audiences with regard to the way they see and read a film. Let's return to *Romance* as an example. There is a scene in the film where the female protagonist has had an argument with her boyfriend, and has decided to pick up and have sex with a stranger she has met in a bar. The two characters are naked and lying in bed, and are clearly about to have intercourse. As the man moves his penis towards her vagina, the woman comes out with a monologue about the ways in which men take sexual advantage of women. The man stops, pulls back his penis and looks (vaguely) thoughtful. The woman then produces a second monologue, this time about how it isn't that simple—that sex is not simply an issue of domination. The man listens to what she has said and, taking her words as a positive signal, moves his penis towards her vagina a second time. Once again, however, she produces a monologue that seems to contradict her previous utterance ('And yet'). The man again withdraws his penis and goes back to looking thoughtful.

Now there are a couple of ways in which this scene (and, because this scene is reasonably representative, the entire film) can be read or responded to. The first response is in terms of sexual excitement. The sight of naked bodies about to engage in sexual intercourse is content normally associated with pornography—and, of course, one of the more obvious social functions of the pornography genre is to engender sexual excitement and pleasure. The second response, more or less diametrically opposite to the first, is to laugh—to treat this scene (and the film as a whole) as ludicrous, pretentious and (unintentionally) comic. After all, the very mechanical back-and-forward movement of the penis, and the incongruous combination of sexual activity and philosophical clichés, means that pathos is in

danger of being supplanted by bathos. But the fact that this film was shown in art cinemas, received very positive reviews from critics and was the work of a 'serious' director probably meant that those two responses were foreclosed—at least for many people in the audiences which saw it.

Genres then, like intertexts, do not provide us with special access to visual reality; rather, they are frames and references that we use to negotiate, edit, evaluate and in a sense read the visual as a series of texts. And the way in which socio-cultural fields and institutions categorise people, places, events and texts in terms of certain genres (often based on or associated with evaluative binaries such as normal/perverted, civilised/barbaric, good/evil, art/pornography) orients and disposes us to see and read the visual world in particular ways.

What is important, in any consideration of how we read the visual, is that as 'readers' we are also 'writers', selecting, editing and framing all that we see. Most of the time this work is unconscious, but even when our seeing is conscious and attentive, we will still make what we see by using the same kinds of techniques (such as selection and omission), and be limited in what we see by factors such as context, habitus and cultural literacy. In our next chapter we extend this inquiry to take into account what we could call the 'prosthetics of seeing'—that is, we consider the relationship between visual apparatuses and technologies, and the types of 'visions' they produce.

conclusion