15 Representations, realism, and history

We paint the remote past, as it were, upon a canvas in our memory, and yet often imagine that we have direct vision of its depths. (William James 1890: 1.643)

It is true that some philosophers have spoken of memory as making us directly acquainted with the past, but all that this comes to is a declaration that they trust their memories. (A.J. Ayer 1976: 135)

Assumptions of direct access do not obviate the need to specify the mechanisms responsible for such feats. (J.D. Bransford, N.S. McCarrell, J.J. Franks, and K.E. Nitsch 1977: 445–6)

15.1 Distributed representation and direct realism
Many Anglophone philosophy students coming to the topic of memory find themselves expected to get excited by old debates about representations and realism. My undergraduate assignment was on the role of images in memory, and required me to discuss whether the concept of memory was a causal concept or not: the underlying issue I was meant to address was the direct or indirect nature of cognitive access to the past. Memory would also crop up in discussions of personal identity, and in strange worries about whether I could know that the world was not created five minutes ago. Disagreement over the nature of memory storage between, say, local and distributed theorists seemed, from these perspectives, to be of minor interest, important only to psychologists: it certainly would not contribute to our understanding of the self. Interdisciplinary contact extended only as far as rude philosophers vainly lamenting conceptual confusion behind empirical assumptions about the trace (Malcolm 1970).

The debate starts roughly like this (Woozley 1949: chs. 2–3; Shoemaker 1972; Locke 1971: chs. 1–4, 8; O’Connor and Carr 1982: ch. 5). Direct realists, hostile to the memory trace, claim that in the act of remembering I am in direct contact with past events. Memory is ‘an immediate knowledge of something past’ (Reid, Essays III.7: 357), or ‘the mind’s awareness of past things themselves’ (Laird 1920: 56). The representationist, in contrast, thinks that an explanation of remembering will include reference to a trace acquired in past experience which somehow ‘represents’ the events remembered. We do know the real past, says such an indirect realist, but only through the mediation of representations which exist in the present.
Neurophilosophical theories of memory, from Descartes to new connectionism, seem at first to fall on the representationist side. According to the Cartesian philosophy of the brain, for instance, traces are left in the fibrous brain substance through which animal spirits flow: remembering is the reconstruction of a patterned flow of spirits relevantly similar to the pattern present at the time of experience. Past experiences are part causes of present remembering: only from, or in, the present events can the past return or become present. Remembering is inference, mediated by traces which represent the past events which partly caused them.

But in fact the relation of distributed models of memory to representationist traditions is far from simple. Both animal spirits and new connectionist models blur alleged boundaries between opposed approaches to memory, and satisfy many of the direct realists’ requirements without giving up on science. The local/distributed distinction turns out to be precisely the point on which the philosophical issues hang.

In chapter 1, I mentioned Rumelhart and Norman’s suggestion (1981: 2) that distributed models might ‘offer an alternative to the “spatial” metaphor of memory storage and retrieval’. They asked (1981: 4): ‘Can the theories developed within the place metaphor of memory be translated into a spatially distributed equivalent?’ Like them, I have no simple answer: but I show that the distributed models can at least resolve or dissolve many specific problems within the long-running debates.

In this chapter I examine what happens when obstructive representations and nasty traces are jettisoned, arguing that direct realism in psychology and in the historiography of philosophy fails, as a positive alternative, to rule out legitimate questions about the memory processes and mechanisms which underlie our attunement to the past. In chapter 16 I set out a new taxonomy of criticisms of memory traces and suggest that, while existing responses leave residual work for the representationist to do, specifically distributed models can do it. Many criticisms of ‘the fiasco of the theory of traces’ (Straus 1962/1966: 99) depend on the mistaken assumption that traces must be local, independent atoms which faithfully store and reproduce the past.

---

1 Many other causal factors are involved: the state of the system at the time of the original experience, the effects of experience in the interim, and the context of reconstruction in the present all influence the fate of the trace.

2 In the case of perception, Gardner (1985: 321–2) suggests that connectionism might reconcile Gibsonian realism and representationism (compare Humphreys and Riddoch 1986; Hatfield 1988b, 1989; O’Brien 1988). The problem in perception is whether we see objects directly, or only through representations partly caused by those objects. The object of perception is usually present, whereas that of memory is, on most views, usually not: direct realism is thus harder to accept for memory. I transfer and adapt arguments from the (extensive) literature on perception to the case of memory.
But I am not only out to confirm for the case of memory ‘how prone anti-representationalists are to attacking imaginary opponents’ (Wright 1993: 1). There is a stronger reason for reassessing this dry literature. The orthodox presentation of indirect realism includes assumptions about the nature of the self whose access to world or past is mediated by representations. Some objections to the trace are misplaced attacks on these implicit pictures of subjectivity. The different form of representationism I defend must find alternative descriptions not only of the trace but also of the self.

The link between memory and self can be relocated by way of two initial responses to the traditional debate. Firstly, I show that the trace theories I have described evade some old notions about representations. Then I query the notion of a bounded inner world of pure subjectivity into which only representational proxies of objects intrude.

Firstly, there is no direct awareness of a trace or an idea from which the subject then indirectly infers the past. Representations were once meant to be immaterial or ontologically ambiguous mental items scanned, in turn, by a non-physical soul. Indirect realism was thus often seen as a form of dualism (Woozley 1949: 21–2, 29–34, 38, 53; Gibson 1979a: 223). When representations are thought of as brain states, it is clear that they are not immediate objects of experience which a subject then consciously puts to use. Some critics, keen to convict Descartes of incoherence, took him to accept direct consciousness of images etched on the pineal gland: but there is no good evidence for this interpretation and ample reason to reject it (Rozemond 1989: 228–43).

But when, secondly, ‘inference’ in perception or memory is unconscious, there can be no incorrigible awareness in the present of a private object (idea, image, sense-datum) from which the past is read off (Shoemaker 1972: 268; Harrison 1976). If the notion of privileged access to transparent mental entities was all the direct realist rejected, there would be no conflict with traces (Pitson 1986; Schwartz 1994: 10–17). ‘Inference’ in distributed models is an immanent computational process, as a system settles into a stable state, with no subject scanning input and itself computing conclusions from pure sensory data. The base worry about a representation acting as a veil or iron curtain, a useless third thing between subject and object, needs to be twisted by dropping, not the representation, but the separate subject. Without a unitary subject behind the traces, there is no principled way to distinguish input from transformed computational representation, data from hypothesis, or ‘the given’ from its supplementation (Schwartz 1994: 116–21, 142).

15.2 Gibson and Gibsonians on remembering

J. J. Gibson (1979a: 254) criticises the concept of 'memory':

Because we are led to separate the present from the past, we find ourselves involved in what I have called the 'muddle of memory'. We think that the past ceases to exist unless it is preserved in memory. We assume that memory is the bridge between the past and the present. We assume that memories accumulate and are stored somewhere; that they are images, or pictures, or representations of the past; or that memory is actually physiological, not mental, consisting of engrams, or traces; or that it actually consists of neural connections, not engrams; that memory is the basis of all learning; that memory is the basis of habit; that memories live on in the unconscious; that heredity is a form of memory; that cultural heredity is another form of memory; that any effect of the past on the present is memory, including hysteresis. If we cannot do any better than this, we should stop using the word.

The rhetoric of this summary dismissal of all theories of memory, apparently on the basis that they differ from each other, is not untypical of Gibsonian ecological psychology. 'Direct realists' have written much less about memory than about perception, and most is only critical. I concentrate on Gibson and his followers, neglecting other positive anti-representationist approaches to memory only through constraints of space, and construct a positive view. Distributed representationists also complain about static traces, logicist cognitive science, and the theoretical over-intellectualising of cognition: perhaps Merlin Donald (1991: 362) is correct to judge that in neo-associationism 'Gibson's neo-Gestalt perceptual ideas were given a neural foundation', or perhaps this cross-classifying of the old distinctions over which controversy raged shows that they had outlived their usefulness (Schwartz 1996: 90).

Gibsonian realism has been incorporated into the history of the new dynamics in cognitive science (van Gelder and Port 1995: 38–9). But the edge is taken off its polemical tang, as if Gibson was recommending only a temporary shift of attention away from 'models of internal mechanisms when the structure of stimulus information remained so poorly understood' (1995: 38). In fact memory research led the way in integrating ecological care for real-world settings into psychology. After a period of debilitating tension between lab-based cognitive psychologists and 'everyday-memory' theorists (Banaji and Crowder 1989; Edwards and Middleton 1990), the combined effects of Neisser's wise unifying work (1982, 1988), of increased attention to autobiographical remembering (Rubin 1986), and of a need to find disciplinary consensus and relevance in controversies over recovered memories have produced more naturalistic research on interactions between rememberer and environment (see Koriat and Goldsmith 1996 for an extensive critical review of recent theories, methods, and metaphors).

Even though complaints about mainstream cognitivism made by early
connectionists were very close to those in some Gibsonian work (Jenkins 1977; Carello et al. 1984), other Gibsonians adopted a messianic tone in calling for the salvation of psychology from representationist errors (Reed 1983, 1986), or for refutations of ‘the traditional conception of memory’ (Wilcox and Katz 1981a: 227). 4 Gibson himself dealt with memory only in passing, with ‘tentative proposals’ to replace ‘the outworn theory of past experience, memory, and mental images’ (1979a: 263): his supporters admit that a Gibsonian approach to remembering needs to be judged with some lenience (Michaels and Carello 1981: 185–6). But selective analysis of Gibson’s later writings, in fact, raises doubt about whether there really is room in his approach for the explanations of long-term memory which the reviled traces were postulated to provide. Like Steen Larsen (1988: 338–9) I argue that an ‘ecological mnemonics’ analogous to Gibson’s view of perception ‘would have to take into account that most of the information that in the past specified states and events in the world is not available in the environment at the time of remembering’ and might thus have persisted internally.

Occluding edges and the perception/memory dichotomy

For Gibson, perceiving organisms are ‘eye-head-brain-body system[s]’ (1979b/1982: 222) in direct contact with objects in their environment. The laws of ecological optics are ‘regularities known to common sense’, learnt simply by observing everyday events, and ‘not by mastering a discipline in school’ like laws of physical optics. Of these ‘laws, rules, or regularities’, one is particularly telling, both for Gibson because it ‘probably cannot be formulated at all in the usual terms of physical optics’ (1979b/1982: 217–19), and for me because of its applicability to memory theory. ‘The discovery of the occluding edge’ (1979a: 189) was, thought Gibson, a radical challenge to inferential paradigms in perception. Contrary to the assumption that you can only see what you see now, an object or a surface can be perceived without being ‘present to the senses’. How is this so?

The occluding edge of an opaque object of sheet screens (hides, conceals, or puts out of sight) part of the background, all of the far side of the object itself, and any other object that is temporarily positioned within the envelope of its solid visual angle. But occlusion is progressive and reversible. The occluding edge can be seen as such and hence to that extent, the hidden as well as the unhiden surfaces can be seen. (Gibson 1979b/1982: 218–19) 5

4 Overall assessments of Gibson (which focus, as he does, on perception) include Richards 1976/7; Fodor and Pylyshyn 1981; Aurell 1984; Morris 1989; Stroll 1989; Hamlyn 1990: ch. 4; and Schwartz 1994: 125–52. For the view that Gibson’s official hostility to processes and mechanisms does not render his work in principle incompatible with them see Marr 1982: 29–31; and Manfredi 1986.

Since an object occluded in this way does not have to be a source of sensory stimulation in order to be perceived, Gibson says, traditional dichotomies between perception and memory and between present and past experience break down. It is, says Gibson, just a relic of old faculty psychology to think that perception is in and of the present, while memory is of the past (1966b/1982). Seeing what is hidden when 'the occluding edge is seen as such' suggests instead that perceiving, remembering, and expecting are all activities with duration, but that the question of what duration each has makes no sense. There is no moment at which a present percept becomes a memory (1979a: 253–5). Indeed, there is no basis to the naive trichotomy between remembrance, perception, and expectancy, between hindsight, sight, and foresight: perception, in a sense, includes all three, for what actually occurs when the present 'hides' past and future in the fact of occluding edges is that 'of course one simply apprehends the environment' (1966a: 276–7; 1975/1982: 396; note the characteristic rhetorical qualifiers).

Gibson's realisation that percepts do not need to be converted into memory traces leads him to claim that past objects and events themselves remain with us as objects of remembering. This encourages both negative polemic and positive trace-less approaches to memory. I look at the two directions in turn.

Memory traces, attunement, and the luxury of long-term memory
Gibson ridicules all non-Gibsonian memory theory, from static engrams to neural connections, because they locate memory information in organisms rather than in the world. Since, for Gibson, percepts are not converted after a momentary stimulus period into traces, no 'short-term memory' system is required. Perception requires no intellectual operations of integration, combination, unification, or comparison of initially discrete sensory images. New perceptions do not need to be compared with existing memories in an act of judgement, because they already contain all the significant information the organism needs (1979a: 221–2, 246–9, 251–3; Gibson and Gibson 1955/1982). Internal differences between organisms are not cognitive differences. The point of the claim that there is no shunting of information from a perceptual faculty to a storage system is not that encoding is already elaborative, new input already feeding the same system in which it will persist: it is that there are no such cognitive systems to hold information in the mind.

In an important paper on temporal order, Gibson discusses Lashley's search for the physiological basis of memory. Noting his rejection of storehouse models and existing trace theories, Gibson picks up an alternative. Lashley had suggested (1950/1988: 63) that 'the learning process must consist of the attunement of the elements of a complex system in such a way that a particular combination or pattern of cells responds more readily than before the experience'. Impressed, Gibson marvels at the novelty of Lashley's idea:
If learning is a kind of resonance in the nervous system, a tuning of the system to certain inputs, then it is not any sort of storage of engrams or depositing of traces... It seems to me an astonishing idea, that learning does not necessarily depend on memory as it has always been conceived. (Gibson 1966b/1982: 172)

This naive passage is significant, for Gibson was about to develop the key concept of 'affordance', describing what it is in environments to which organisms are attuned (the manuscript 'Notes on Affordances' dates from August 1967). But the dichotomy set up here by Gibson is a false one. Trace theories and resonance theories are not incompatible. Hartley's vibration theory of memory (chapter 13 above) is a resonance theory if any is: and yet there is clearly a use for the term 'trace' for Hartley, even though it does not name an entity permanently and passively stored at a fixed memory address. The same goes for new connectionist models, which talk of resonance in more detail, and yet continue to use the concept of trace outside the crude localist spatial metaphor: traces are the means by which the system attunes itself to certain inputs. Gibson is blind to the historical existence of anything other than faithful local reproductive computer-like models of memory traces.

But Gibson, trying to bypass 'the muddle of memory', then claims that an explanatory theory of information pickup will not refer to memory. Lashley's 'hypothesis of tuning or resonance... suggests a surprising possibility - that learning does not depend on memory at all, at least not on the re-arousal of traces or the remembering of the past' (1966a: 275). The state of a system is altered or 'sensitized' when attuned to information of a certain sort. But, says Gibson, this learning, this 'altered state need not be thought of as depending on a memory, an image, an engram, or a trace. An image of the past, if experienced at all, would be only an incidental symptom of the altered state' (1979a: 254). As well as the usual assimilation of quite distinct targets here, Gibson denies that there can be further questions about how attunement occurs, about what its mechanisms might be. The danger of obscurantism in this form of direct realism is marked.

Images, said Gibson, are only incidental accompaniments of attunement. Clearly he is not seeking a form of representation which does not rely on resemblance, as indirect realists who reject the necessity for images might. Instead he is disclaiming any need to explain phenomena of long-term memory. The rejection of traditional theories of memory runs deep: there is to be no account of 'recollection' understood as, for instance, voluntary recall of events long in the personal past. The only phenomena admitted will be forms of habit memory where an organism is sensitised to new information and 'differences are noticed that were previously not noticed. Features become distinctive that were formerly vague' (1979a: 254). One can accept the importance of these phenomena, especially in ethological psychology: but that is not to say that they are all that an approach to learning and memory needs to explain.
Yet this is what Gibson intends: ‘what we call memory and recognition is often only a special case of invariant-detection. Recollection is then a sort of human luxury, an incidental ability to contemplate the past’ (1966b/1982: 178). This is strange science for one concerned with biological plausibility, for the human ability to contemplate the past can hardly have had an incidental evolutionary importance. Gibson refers to recollection again only to say that it is not explained by ‘the traditional hypothesis of mental imagery’ (1979a: 256). After discussing the perception of occluded surfaces again, he notes that ‘one kind of remembering’ is ‘an awareness of surfaces that have ceased to exist or events that will not recur, such as items in the story of one’s own life. There is no point of observation at which such an item will come into sight’ (1979a: 255). The information gained in such activity is ‘further detached from stimulation’ than is perceptual information: and other than the polemical reminder that ‘[this] is still an activity of the system, not an appearance in the theater of consciousness’ (1979a: 256), that is all we are told. This is not a replacement theory of memory but a denial of its relevance or importance. Gibson felt little need to extend treatment of his proprietary domain, perception, to cover remembering: perceiving, for him, was ‘the simplest and best kind of knowing’ (1979a: 263).

Gibson’s followers likewise displace human long-term memory as an explanandum. Turvey and Shaw (1979) set out to prove that Gibson’s work provides better directions ‘for understanding memory’ than does ‘the legacy of the past five centuries’. Yet, remarkably, in a 55-page paper they refer to memory only in the three concluding pages, which propose only a view of memory ‘as knowledge that persists by analogical extension (generalization) from earlier to later situations’ (1979: 219)!6 Michaels and Carello, also dispensing with any ‘concept of memory’ which requires ‘the storage or retrieval of information’, give no alternative other than a plea for relating ontogeny to phylogeny by collapsing individual learning and memory into evolutionary adaptation. They complain that traditional sciences ‘never view the evolutionary concept of experience as the amassing of memories while they rarely view the consequences of an animal’s personal history as anything but the amassing of

6 Turvey and Shaw argue that memory is a property of an ecosystem, and not of an organism. They are thus forced vigorously to deny that the causal processes which support memory are relevant to an understanding of ‘the epistemic act of remembering as such’. These claims, however, are supported only with quasi-formal ‘postulates’ which already assume the attribution of memory only to ecosystems (1979: 217–19). But even before pushing the criticism that the neo-Gibsonian formalisms are misleading in failing significantly to constrain possibilities (Cutting 1982: 211–13), one may doubt the force of Turvey and Shaw’s rejection of memory traces. They deny that traces left by experience could be things which ‘enter into a recipe whose product is perception’, but they accept that experience, which is ‘preparatory to perceiving’, ‘attunes or sensitizes perceptual systems to the information that specifies affordances’ (1979: 216–17). The dichotomy between trace theory and attunement or sensitisation is, again, unfounded.
CONNECTIONISM AND THE PHILOSOPHY OF MEMORY

memories' (1981: 77). This is an astonishing suggestion: cultural change and its effects on individual learning are to be understood in exactly the same terms as the effects of evolutionary history on the individual's adaptation (compare Schacter 1982: 105–47; and Otis 1994 on 'organic memory'). Both evolution and individual learning produce 'a new animal that is better able to cope with its environment'. Experience, we are told, actually leads to a new machine: 'the consequence of personal experience is not that the old animal has new knowledge, but that it is a new animal that knows better' (Michaels and Carello 1981: 78; emphasis is in the original). This is a strange way to argue either for biological plausibility in psychology or for the disunity and discontinuity of personal identity. Any more detailed attempt to explain the phenomena of long-term memory would have to start looking for mechanisms: but the hostility of Gibsonians to psychophysiological explanation prevents them from taking such phenomena seriously. If traces and attunement are mutually exclusive, then denying traces looks sensible. But there is no good reason to accept the antecedent.

Temporal order and the specification of the past

The Gibsonian programme encourages reconsideration of temporal structure in the explanation of remembering. Asserting that we are aware of persistence in change, that we do not have to construct stability out of a sensory flux, Gibson says that perception is of sequences, not of discrete images. The stability is out there, in the invariants of the optical array (1975/1982; 1979a: 221–2, 248–9). He rejects the Humean picture of atomic sensations succeeding one another like snapshots without real connections (1979a: 250). Sequential and temporal structures are just as real as spatial structures.

The representational theory of memory, in contrast, 'fits harmoniously with the common sense, Western idea of time', which is an unfortunate inheritance of philosophical tradition (Wilcox and Katz 1981a: 233). The supposed need for traces stems from seeing time as 'linear and absolute', and allowing the past no independent ontological status. Stored mental representations are then hypothesised to keep the past in mind. Gibson's scepticism (1966b/1982) about cognitive experiments which rely on the presentation of successions of static patterns is generalised to an attack on the need to bridge alleged temporal gulfs between discrete units of time.

This argument is also in the philosophical literature. John Laird (1920: 46), responding to the attempt to join temporal gaps with traces, claims that all perceptions are of 'an enduring slab of time or a stretch of change'. Such a stretch has an objective order of earlier and later embedded in it. Past events have their determinate place in a series, and so the past 'is neither trackless nor unknown' (1920: 59). Laird asserts that this is enough to show there to be no contradiction in the notion of direct apprehension of the past. Wilcox and Katz in similar vein
envisage time as ordered sequential structure, of which past events are part. Sequential structure, as in a melody, is temporally constrained: before and after, earlier and later, are real distinctions within it. A melody will not be perceived as that melody if its sequential elements are disordered (1981a: 235; compare Earle 1956/7: 22–6; Jenkins 1977: 428; Arcaya 1989: 103–4, 1992). Whole stretches of temporal structure can be apprehended in the present without the need (impossibly) to jump to other instants.

There's ample reason to accept that, for systems like us, time comes in slabs or stretches rather than discrete and unconnected instants. But questions remain about the effects of a rememberer's specific location within such a temporal structure. Even Gibsonians need some account of individual differences in the effects of learning history and experience, and in memory capacities and abilities, if stretches of time are 'simply' directly accessible (Sanders 1985: 512–5; for controversy over individual difference in perception see Heil 1979, 1981, against Reed and Jones 1981, and Wilcox and Katz 1981b). Representationist cognitive theories are good at this, for background memory representations acquired in individual experience have causal effects as context on the nature of particular acts of remembering. The present environment alone fails uniquely to specify the past for any individual. Remembering is underdetermined by either cue or engram alone: differences in past experience affect it through the effects of experience on internal states. Remembering, produced by the interaction of meagre stimulus and internal context, if accurate, is so contingently (Rock 1991; Searle 1991).

I argue in chapter 16 that connectionist theories of memory are not only compatible with, but require careful stress on information about the past being available in the present. Sometimes, indeed, retrieval cues must be highly specific to elicit remembering: memory is not simply 'an activated picture of a past event' (Schacter 1996: 60–4, 71). But present hints and cues do not work on air: 'a neural network combines information in the present environment with patterns that have been stored in the past, and the resulting mixture of the two is what the network remembers' (Schacter 1996: 71). Gibsonians deny that this mixing occurs in the brain or mind, rejecting arguments from the poverty of the stimulus (Michaels and Carello 1981: ch. 1), and claim that the past is uniquely specified in enduring information which somehow remains in the environment. Explaining how this could be so is the central task in the application of Gibsonian theory to memory. Yet existing attempts produce alternatives only to static localist storehouse models of memory, and are, despite their proponents' wishes, compatible with distributed models.

To take one example, Bransford and colleagues, sensitive to Gibson's thought, rejected the 'searching for traces conceptualization' (Bransford et al. 1977: 432). Yet, invoking Bartlett, they suggested instead only a 're-creation
metaphor of remembering' (1977: 449–56) not unlike dynamic connectionism. Past experience ‘sets the stage’ for later experience through the education or attunement of an organism, without being faithfully stored in an internal system (1977: 434–9). This stage-setting can be thought of as an imposition through attunement of constraints on creative and (re)constructive processes (1977: 441–3). Perception and recognition are said to be direct given the attunement (1977: 435). One can accept the need for ‘an account of the global level of “attunement” set by the context of the situation’, a level which ‘will affect the nature of the experiences afforded by potential cues’ (1977: 443). But why does one individual reconstruct the past in this way, and another quite differently? Because of their different existing levels of attunement (1977: 450). Here hypotheses about what it is that grounds or underpins the differing attunements which set constraints on resonance are hard to rule out of court on conceptual grounds.

To repeat, there is no reason to think the opposition between storage and reconstruction on which Bransford et al. rely (1977: 453) to be a true dichotomy. Of course the aspects of the world to which an organism is currently attuned matter for remembering: internal states do not just reactivate in a vacuum. Yet caricatures of storage as faithful reproduction continually encourage anti-theory rhetoric (Arcaya 1989: 104; Ben-Zeev 1986). Direct realists, from Reid on, are drawn to reject all analysis and explanation of the phenomena of remembering. Laird (1920: 59) complained that ‘it is quite unreasonable to be dissatisfied with the analysis and description of apprehension as we find it. We can perceive the present and recollect the past; and we are not required to explain the inexplicable.’ Joel Michell (1988: 247) quotes this passage to deny that representationist theories can explain ‘the emergence of cognition’. He agrees that ‘the occurrence of cognition will have its necessary and sufficient physical and physiological conditions’, but denies that specifying them explains how cognition is ‘generated by these conditions’ (Michell 1988: 247). Apart from the fact that many representationists do not believe that there are necessary physical conditions for cognition, this blanket rejection of explanation is unnecessary caution, disallowing common ground. Michell is deploying a powerful critique of cognitivism by John Maze (1983): but Maze, in contrast, is willing to wonder on the nature of brain traces. Maze marvels at the intricacy of the neural mechanisms underpinning cognition, mechanisms which must operate in a variable ‘context-related’ way, without there being ‘one separate neural trace for each proposition known’ (Maze 1983: 91). The possibility of reconciling a distributed connectionist approach to memory with direct realism could not be clearer.7

7 I have not touched on the problems which direct realists face in accounting for error (compare Cutting 1982: 209–11). Gibson himself allowed for rare perceptual errors, when for example ‘a thing may not look like what it is’ (1979a: 143), but his followers sometimes deny the intelligibility of misperceiving and misremembering. Turvey and
15.3 Gibsonian history

Polemic against representations is also prominent in interpretations of early modern philosophy. Descartes' indirect realism, complains Marjorie Grene (1985: 211) is responsible for 'disastrous, science-fiction philosophising'. Rejecting Descartes' internal processing of visual information, Grene (1985: 205-12) invokes the full Gibsonian story about the existence of information in the ambient light which uniquely specifies what is seen. As I suggested in section 15.2, it is not easy to apply this picture to remembering. Grene recommends as 'a more reasonable beginning' the attempt, with Gibson or Merleau-Ponty, 'to put intellectual activities back into their place in our natures as living things, and thereby also to put meaning back into the natural world', a world in which, 'though in a complicated and often messy way, things make sense' (Grene 1985: 212, 208-9). It is hard to disagree with such wisdom! But obviously I deny that the attainment of genuinely 'full-bodied' perspectives on psychology requires exposure of a 'crippling Cartesian heritage' (Grene 1985: 199, 95).

Grene addresses conceptions of time, arguing with Gibson that we are sensitive to processes, 'not snapshot-like presentations' (1985: 202). As well as bringing temporal context into the perceptual process, Grene insists on the importance of personal/historical contexts, and on the activity of the embodied organism with its needs. This is all laudable: but there is no reason why a context-sensitive distributed representationism cannot accommodate it, and no reason but for fear of mechanism to deny even that Descartes' model of memory allows it (chapter 3 above). Grene moves easily from care for contexts to the rejection of mechanism. 8 But mechanism is not refuted by pointing out

Shaw (1979: 213) note that 'to say that perception it is direct, rather than indirect, is to say that it necessarily provides information about how things are (in reference to the effectiveness of a given animal) and not merely about how they appear'. The fact that indirect realism fails to guarantee the veridicality of perception is, for them, an argument against it (1979: 178, 182 n. 2, 214)! On memory, Wilcox and Katz (1981a: 235–7) suggest that, since memory is just 'the apprehension of elements that are part of a sequential structure', the notion of error must be 'relativistic', and has sense only as 'the perception of an element in a sequential structure that has changed as we discover more about the world'. Denying error in remembering is unacceptable to many who are otherwise sympathetic to direct realism, for on most views the possibility of error is the price of realism. For a complaint that Gibsonian observer-dependence in perception buys its directness at the cost of a relativistic loss of the independence of perceiver and environment, breaking down a needed 'distinction between objects in themselves and objects as known', see Yolton 1987: 329-30; compare Yolton 1996: 21-30.

8 This is evident in an earlier essay on Hobbes (Grene 1976), in which to Hobbes' materialism and representationism are quickly assimilated all the 'barbarisms' and 'computer-ridden speculations' of scientistic modern philosophy, from the loss of God and the denial of significance in the universe, through ignorance of the 'higher realities' of human culture, to reductionism, determinism, and the rejection of teleology, of 'oughts', laws, reason, and choice. One might wonder why, again, it is the anti-reductionist who assimilates many quite different things.
that organisms are not passive, for internal causes can be explained mechanistically just as can external stimuli. Neither is it refuted by pointing to the importance of the personal past (Grene 1985: 202–3), for the personal past has its effects by way of superposed memory traces which interact, blend, and form again according to the present situation.

Responding further to historians’ uses of Descartes as exemplar and originary sources of persistent representationist errors allows me to sketch a more naturalistic reading of historical indirect realism. Like John Yolton, I hope that history ‘may even offer some resolutions to questions about representation and realism’ (1990b: 516). Yolton, attuned to Gibson’s work for some time (Yolton 1968/9), has led a different reinterpretation by revivifying Arnauld’s act theory of representative perception, which drops representational entities in favour of a significatory relation between perceiver and world (Yolton 1984a: ch. 1, 1987, 1990a). I conclude with a brief response.

These historians focus almost exclusively on perception, with theories of memory, let alone those postulating distributed representations, barely featuring. It is hard to interpret early modern philosophers as accepting the action at a temporal distance which non-mechanistic theories of memory require. Yolton (1987: 327) notes that ‘the influence of memory’ is one of the numerous details not covered by writers in the ‘way of ideas’ tradition. Yet some of them, I have shown, did have neurophilosophical theories of memory.

**Representation and world**

For Edward Reed, Gibson’s biographer, Gibson offers ‘the first truly new theory’ in psychology ‘for the last 400 years’ (Reed 1988: 2), allowing escape from ‘Cartesian themata’ which are disastrous ‘intellectual straitjackets’ (Reed 1990: 101–2, 1989). Recently Reed questions the assumption or hope that a self can move or control its own physiological processes. As I showed in parts I and II, the idea of the mind ‘slyly affecting’ bodily movements by way of animal spirits leads to the Cartesian moral injunctions to keep the will in line, to maintain inner discipline by ‘conquering the passions’ (Reed 1990: 119, 112–13). The will is meant to be limitless but is actually bound within the body: since physical objects are transformed into representations which do not resemble their causes, the will acts only on the internal proxies, not on the external objects which common sense takes to be the objects of volition. We are unaware of mental action on the animal spirits, so ‘much or all of volition must be unconscious’ (1990: 109, 116–17). Where Reed thinks this the absurd legacy of Descartes’ refusal to understand volition as directed at the environment, from

---

9 Compare Yolton 1987: 330. Yolton 1996 impressively integrates early modern and contemporary accounts of perception and representation: I regret being unable to use it more fully. It is not clear how memory would fit Yolton’s schema. My thanks to John Yolton for helpful correspondence and discussion about this material.
REPRESENTATIONS, REALISM, AND HISTORY

which only Gibson offers escape, I have taken it to be a roughly accurate approach to the plight of cognition.

Reed’s rejection of representations relies on a slippage between two formulations of Descartes’ position. One is that ‘all existing thoughts are consequences of the motions of the brain’; the other that ‘all awarenesses are awarenesses of brain states’ (Reed 1982: 733). But these are not equivalent: the former can be accepted while the latter is denied. The hardest materialist reductionist claims only that awarenesses are brain states, not that they are of brain states. Descartes’ insistence that ideas do not resemble their objects, that the world of appearances differs from the physical world, does not entail that we only ever see our own brains, but does drive his rejection of direct realism (Gaukroger 1990: 37–9). But Descartes explicitly allows for a number of different relations of representation which do not require resemblance, and which together will explain why the appearances are as they are.

Descartes’s distinction between representation and resemblance, as Larmore observes (1980: 13–16), derives from his physiology: the relation between patterns of animal spirits motions and perceptible qualities of objects cannot, obviously, be simply pictorial. Representation by resemblance would be incompatible with science as well as superfluous (Meyering 1989: 82–4, 105–6; Hatfield 1990: 52–3). Descartes outlines a general notion of representation, by which words represent what they signify, laughter and tears represent joy and sadness, two-dimensional engravings represent three-dimensional objects, and brain patterns represent objects and their properties.

Nancy Maull, in turn, shows that Descartes’ account of distance perception explains why we are unaware of the unconscious processes which underpin it (Maull 1980: 27–30, 30–4). Maull takes natural geometry to be algorithms for generating judgements about distant objects, objects which are indirectly accessible ‘because of the intervention of geometry or reasoning about [visual] cues’. This ‘elaborate optical plumbing . . . does not imply that the soul is aware of a pineal pattern’ (1980: 30, 34; compare on distance perception Lennon 1980; and Gaukroger 1990: 18–26). Ann Mackenzie, summarising passages on different ‘representational situations’, identifies three main elements of the theory of representation. Brain events, specifically corporeal ideas, ‘are viewed as vehicles of representation; physical objects and/or their properties are viewed as the objects of representation; and the whole causal nexus which links registration events in the brain to objects and properties represented constitutes the background system by virtue of which the vehicles represent their objects to the mind’ (Mackenzie 1989: 179). This interpretation allows for (indirect, fallibilist)

---

10 Maull diagnoses scholarly failure to attend to ‘the peculiarities of Descartes’ version of representative perception’ as partly due to ‘an orthodoxy of misplaced emphasis on Descartes’ more “philosophical” texts’ (1980: 40, 35). Compare chapter 3 above.
realism, and takes seriously the contextual role of the causal background. It still requires, as Mackenzie acknowledges, an unexplained external user of the representational system in the form of the mind (Mackenzie 1989: 190).

In another recent invocation of Gibson, Emily Grosholz draws on his dichotomy between (mere) physical causation and ‘a revised and extended notion of causality appropriate to biology’ to accuse Descartes of confusion because animal spirits are described at both physiological and psychological levels (Grosholz 1991: 129–31).\textsuperscript{11} The distinction between, on the one hand, physical-causal relations between perceiver and world and, on the other, some quite different kind of relation is developed further by Yolton. But where Grosholz complains that Descartes failed to keep them separate, Yolton sees in Descartes at least hints of a distinction between different relations.

The standard history of philosophy, on which Descartes ‘invented the mind’ and inserted a veil of ontologically ambiguous ideas between knower and known (Rorty 1980: ch. 1), is questioned by Yolton who also examines the source of this standard interpretation in Reid (Yolton 1984a: ch. 11). In contrast Yolton argues that evidence is slim for the allegedly widespread early modern commitment to the view that perceivers are only ever directly aware of reified intermediary ‘third things’; slim too for a consequent sudden concern with epistemological scepticism because of the gap thus opened between subject and reality (Yolton 1981, 1990a: 67–70; compare Jolley 1989: 1–11; Hatfield 1990: 46–60).

Instead of the naturalistic reinterpretation of indirect realism offered by historians sympathetic to new connectionism such as Hatfield and Meyering, with which I have been working throughout these studies, Yolton and others seek to break the impasse in modern debates with a form of direct realism which yet retains a role for both cognitive and causal processes. The key is in Arnauld’s development of Descartes, whereby ‘all our perceptions are essentially representative modalities’ (Arnauld 1683/1990: 66). Where Malebranche took ideas (in the mind of God) to be the objects of perception, Arnauld understood them as acts of perception, and thus not intermediary objects at all. An idea just is a perception or, better, a perceiving (on the Arnauld/Malebranche dispute see Laird 1920: ch. 1; Cook 1974; Radner 1976; Wahl 1988). So Arnauld was a direct realist (Nadler 1989; Cook 1991), who saw that when the relation of representation holds, not between an idea-object and a physical object, but between an act of

\textsuperscript{11} This kind of mistake (‘confounding levels of description’) allegedly also permeates ‘contemporary materialist accounts in epistemology and cognitive psychology’. Grosholz’s brief invocation (1991: 129, 131 n. 21) of Gibson’s work, whence she draws ‘some of [her] theoretical orientation’, supports the negative claim that a pattern with cognitive significance ‘cannot be physically-causally transmitted’ but ‘must be recognized, interpreted, cognitively grasped, by a consciousness’. See on this chapter 3 above, and compare Gaukroger 1990: 39.
perceiving and a physical object, ‘the indirectness traditionally associated with representative theories disappears’ (Yolton 1991b: 111). Since ideas are acts of perceiving, modifications of mind, ‘perception’ can be used to designate a perception-idea in relation to the mind which they modify, and ‘idea’ the same perception-idea in relation to a physical object as it is in the mind (Yolton 1984a: 61–2).

Yolton traces Arnauld’s direct but representative realism back to Descartes (1984a: ch. 1, 1987: 323–5, 1996: 183–214; compare O’Neil 1974). I discuss briefly just two strands of this reinterpretation: the account of natural signs; and the distinction between causal and cognitive relations. In each case I merely query whether this interpretation is incompatible with naturalism and causal theories of perception and cognition.

Descartes hints that natural signs are instituted by nature or God to occasion perceptual responses. Without internalised copies of objects, a natural geometry makes us respond in a particular way to a particular sign, ‘even if the sign contained nothing in itself which is similar to this sensation’ (Le Monde 1, AT xi.4, CSM 1.81). Yolton’s great insight is to see that ideas themselves are not signs. Ideas are merely ‘the interpretations of, or cognitive responses to’ the signs, where the signs are physical (physiological) motions in nerves and brain (Yolton 1990a: 62; compare 1984a: 22–31). So we do not require a self separate from its ideas to read information about objects off from its idea-signs. This is why Descartes insists that our responses to all perceptible differences in objects are occasioned by differences in the motions in the brain: these motions, not ambiguous ideas, do the signifying.

But Yolton keenly distinguishes his view from causal or information-theoretic accounts, by which corporeal motions encode information about the external objects which caused them. Descartes rejects ‘any causal relation between the physical activity of objects on our sense and the perceptual ideas in our minds’ (Yolton 1984a: 18): the point of using natural-sign language in ‘describing the act of perceptual cognition . . . is to avoid the language of causation’ (Gaukroger 1990: 24).

I focus on this incompatibilist account of causal and cognitive relations below: but even within the natural-sign doctrine there are reasons for resisting it. The ‘teachings of nature’ formed in childhood are habitual judgements made on the basis of natural signs (sixth set of Replies, AT vii.438–9, CSM 11.295–6), instituted by God for the preservation of the body, not the

12 Yet Yolton resists the excessive claims of Gibsonian realism, questioning the complete rejection of physiology and the wish to dispense entirely with psychological processes (1987: 327–30). No realism can be so direct, Yolton acknowledges, as to ‘escape from the cognitive, psychological, and representative nature of perceptual awareness’. Unless the perceiver’s mind is, impossibly, just out there physically wandering among objects to seize their being, there must be cognitive terms in an account of perception (1987: 329–30).
perception of truths (Fourth Meditation, AT vii.82, CSM 11.56–7; compare Hatfield 1986). The judgements of Descartes’ third grade of sense are unnoticed responses. But now, to explain the possibility of error in these responses, might it not be that we have to look at deviant causal processes between object, sense organs, and brain? Where Descartes, reluctant to allow non-conscious cognitive processing, simply spoke of automatic responses instituted by nature, Malebranche and later cognitive theorists would fill out different ways in which these irresistible God-given responses occur (Meyering 1989: 89–105). 13 What determines differences in perceptual or cognitive response to natural signs are differences in the signs themselves, caused by different physical-causal relations between the perceiver and the physical object.

Indeed Yolton’s own account is sometimes often couched in terms which look like those of a causal theory. Awareness is ‘a result of the cognitive activity of the mind working in conjunction with (even attending to) physiological processes in the brain’ (Yolton 1984a: 39): but perceptual discriminations are made ‘on the basis of sensations felt by the perceiver’, sensations which are ‘a response to or an interpretation of natural signs’ (1984a: 39, 1990a: 62). Elsewhere the nature of the cognitive reaction to physical events is said to be ‘translatory’ (1987: 325). How does this interpretation or translation of natural signs differ from the ‘reading off’ of properties of the world which Yolton officially rejects (1984a: 26)? When he does accept that ‘the mind reads the physical motions [in nerves or brain], as it does the tears and smiles of a face’ (1984a: 30), the only distance kept between this view and a naturalistic causal theory is the negative point that this reading, translation, or interpretation is not causal. To this I now turn.

Although Yolton on occasion wonders if, in the psychology which Descartes outlines ‘to link cognition with physiology’ (1984a: 21), causal and cognitive relations between perceiver and object could go together, 14 he generally maintains the incompatibilist view that ‘for Descartes there is no causal relation in perception between physical objects and ideas’ (1987: 325). 15 The interaction is

---

13 For Meyering, Descartes’ ideas on natural geometry and natural signs suggest that the cues ordained by nature to occasion responses in us are being automatically ‘decoded and interpreted’ in subcognitive information processing (1989: 85–6). Even Malebranche’s God must obey the fixed laws of mind/body union: as Meyering says (1989: 104), ‘only changes that are physiologically imprinted upon the senses can occasion God’s incredible production of exact psychic responses’.

14 Some possibilities Yolton canvasses are that causal and physiological relations, while being ‘inadequate for cognition’ (1990a: 62–3), are supplemented by (1987: 323), or work in tandem with (1990a: 62), or even are (1984a: 39) significatory, cognitive, or semantic relations.

15 Ideas are not causal effects of motions (1984a: 19); Descartes rejects ‘causation between brain and cognitive activity’ (1984a: 21), and tries to preserve an ‘interaction between body and mind which is not causal, or which is more than causal’ (1990a: 62); ‘motion in my body does not cause but signifies my sensations’ (1990a: 62). Yolton now suggests (1996: conclusion) that, after his correspondence with Princess Elizabeth, Descartes replaced his
not causal, but 'significatory'. There is a connection between mind and world, but it is 'precisely that which is proper to cognition: significatory, not resemblance or causation' (1984a: 30, see also 1996: 183–214). Yolton acknowledges that Descartes' brief hints at a 'doctrine of signification' are 'far from clear' (1984a: 30–1, 19).16 Without here going into Yolton's positive development of such a doctrine,17 I want to note how strange it is to deny that minds as well as brains enter into causal relations, especially in the case of memory. Creative or reconstructive cognitive remembering, even on Yolton's view, is an interpretation of the physical traces which stand in causal relations to past events. The language of interpretation and signification reminds us that the trace itself cannot uniquely determine the remembering, that the current cognitive context counts: but, that said, it is hard to see what is lost in saying that the trace is a cause (among others) of the cognitive episode. What explanation could there be of the way memories blend, mix, and interfere if the past is directly signified?

Yolton accurately pinpoints the problems of causal interaction over a substantial distinction between mind and body as the residual sticking-point for early modern causal theories (Yolton 1984a: 22, 1987: 326). However, reinterpreting history for philosophical use, an alternative to rejecting the causal explanation of world/representation relations is to seek dis/solution of the problems of a 'third thing' by rejecting the mental subject behind the traces.

16 For the related view that a theory of action should be not a causal theory but an account of the 'significance of action' see Yolton 1984b: 145–50. This strong incompatibilism between the causal and the significatory seems surprising for one who had previously complained at the 'curious reluctance' of recent action theorists 'to employ the notion of cause' (1966: 17): but in fact the mental causes to which Yolton then referred were incompatible with physical causation (1966: 22–4).

17 For responses to Yolton's account of the significatory relation see Costa 1983; Michael and Michael 1985; Ayers 1986; Matthews 1986; Cook 1987; Kelly 1987.