

Fodor, connectionism, and cognitive discipline

When . . . vibrations occur to the mind in their natural order, and retain their original, natural rational associations and the mind can retard or accelerate their progress, to a certain degree, at pleasure, and can deliberately consider them and compare the past with the present, when of course, memory holds her seat, connecting consciousness, as Hartley terms it, is unimpaired and the associationism of truth and decency and propriety, and virtue remain uninjured; in short when the mind can regulate properly all its operations, it is then a sound and rational state: but in proportion as the reverse takes place, in such proportion is it in a state of unsoundness and insanity. (Thomas Arnold 1786: 284, in Hoeldtke 1967: 49)

11.1 Fodor, connectionism, and associationism

Jerry Fodor follows up his complaint that cognitive science is 'pathetically out of contact with its own history' (1985a: 8) by employing history in support of his own language of thought hypothesis and representational theory of mind. In *The Modularity of Mind* (1983: 14–36) he argues against historical associationism and in favour of a 'vertical' faculty psychology inspired by the phrenologist Gall. The polemical paper 'Connectionism and Cognitive Architecture' (Fodor and Pylyshyn 1988) returns to the attack on associationism. Attention to old shortcomings of associationism undermines new connectionism, a troublesome alternative to the language of thought.

In this 'enjoyable, but essentially digressive, interlude of Connectionist bashing' (Fodor 1990: ix), Fodor and Pylyshyn present arguments against both classical empiricist association of ideas and modern behaviourism which, they suggest, were decisive in their respective historical contexts, and ought to be decisive now (1988: 6). Even connectionism's vaunted neural fidelity is unfortunate, only encouraging old errors: 'The idea that the brain is a neural network motivates the revival of a largely discredited Associationist psychology' (1988: 63).

Where Fodor and Pylyshyn lament associationism's links with neuroscience, others complain at their absence. Finding little concern with neural mechanisms of memory before the late nineteenth century, Schacter and Tulving (1994: 4) explain the lack of neural influence on cognitive theory thus: 'there was little understanding of the brain at the time . . . the doctrine of associationism held almost universal sway over philosophical and psychological thinking about memory, rendering any kind of physiologizing

superfluous'. So the early modern neurophilosophy I describe in this book is a tradition neglected by some and rejected by others. For the moment, I accept Fodor's connections between old and new forms of associationism, and examine his arguments against both. How convincing is the attempt to blacken connectionism by chaining it to a tainted past?

Fodor is unusual in stressing the implication of weighty cultural and political issues in cognitive science (1985c/1990: 206). He does not deny the links between technical debates about the adequacy of particular theories and broader questions about the nature of psychological explananda and the relative importance of order and disorder in cognition. Reasonably enough as a practising philosopher of psychology, Fodor defends himself against carping doubts about scholarly accuracy with regular disclaimers that he is only 'making up some history', a justifiable practice if, as he says of cognitive science, 'we can see something about where we are and where we're up to if you provide a little backlight' (1985a: 4).¹ So I track this useful target through his entertaining corpus. My treatment of his complex philosophical agenda is sparse (see Loewer and Rey 1991 for an introduction): in this context I am interested less in refuting Fodor's arguments than in questioning his starting-points. I sketch his historical contentions, before showing how they fit with his views on broader issues about discipline and argument.

The problem with associationism, in Fodor's eyes, is that it cannot explain the parallel between mental processes and *arguments*. According to Fodor, what makes thinking worthwhile 'is that trains of thoughts should be generated by mechanisms that are generally truth-preserving' in the same way as 'argument is generally truth-preserving' (1985b/1991: 40–1). His own representational theory of mind (RTM) is intended to provide 'a theory of mental processes that succeeds where associationism (to say nothing of behaviourism) abjectly failed; a theory which explains how there could regularly be non-arbitrary content relations among causally related thoughts' (1985b/1991: 42).

Since we now know that it is possible to connect semantic properties of symbols with their causal properties in computers, we can 'abandon associationism and combine RTM with the "computer metaphor"' by showing 'how to connect semantical with causal properties for thoughts' (1985b/1991: 41; compare 1987: 10–21). Fodor's RTM as a theory of mind is explicitly based on a theory of rational mental processes, of inferring, judging, and the like (see Clark 1988 for an excellent critical account). Rejecting the connectionist claim that we must examine 'lower levels' of the mind/brain to understand higher

1 Fodor acknowledges that his use of history is 'no doubt, all pretty loose – a matter less of demonstrative arguments than of elective affinities' (1983: 36). This sensible strategy, based on building up considerations and resemblances rather than on decisive argument, is notable only because it conflicts with his official view that only objectively rational argument deserves assent.

levels of function and processing, Fodor and Pylyshyn think it sufficient to argue that suggested microstructural to macrolevel correspondences or isomorphisms are unlikely to hold when 'the cognitive process under investigation is as abstract as reasoning' (1988: 63).² So proponents of RTM take reasoning as the primary explanandum for a cognitive theory. Are those relatively ordered psychological processes in which truth-preserving mechanisms do seem to operate a suitable model for the rest?

It is not surprising that theorists who privilege inference and reasoning find associationism unsatisfactory. Fodor's hero in the history of philosophy is Kant, who, he tells us, 'let the cat out of the bag' by pointing out what associationists had hardly noticed, that 'thought – like argument – involves judging and inferring' (1985b/1991: 41).³ The leading arguments paraded against connectionism as a cognitive theory follow.⁴ The regularity of trains of thought which, like steps in argument, preserve inferential relations is explicable only by attributing to thought the same combinatorial constituent structure as exists in grammatical sentences of language. The combinatorial structure of sentences explains the systematicity of linguistic capacities: only similar constituent structure in thought could explain the systematicity of cognitive capacities (Fodor 1987: 151; Fodor and Pylyshyn 1988: 37–41). If thoughts have constituent structure, there must exist mental representations or symbols, internally structured and capable of arbitrary recombination with other representations, in a language-like internal representational medium: a language of thought.⁵ The

- 2 Compare another attack on associationism (Fodor 1986: 319), where the problem is that 'the notion of the form of an inference can get no grip in an associative net' because 'the nodes in a network have no form' and thus cannot distinguish between different but logically equivalent formulae on the basis of, say, the order of their parts, where humans can do so.
- 3 Arguing on these grounds that associationism 'failed to produce a credible mechanism for thinking', Fodor goes on: 'no wonder everybody gave up and turned into a behaviourist'. But this really is C-minus history. Kant failed to prevent associationism's dominance through the nineteenth century as a cognitive theory (on German associationism after Kant see Hatfield 1990: chs. 4–5). Fodor is aware, when convenient, of the survival of associationism in the face of rationalist criticism: when seeking causes of the demise of faculty psychology after Gall, he happily blames the continuing 'promise of an associationistic theory of mind' (1983: 27).
- 4 I focus on the case for the systematicity of thought, Fodor and Pylyshyn's central argument: they also argue for the productivity of thought and the compositionality of representations (1988: 33–7, 41–6).
- 5 Officially, Fodor wants to remain neutral on the question of just how often the kinds of regularity and systematicity evident in linguistic capacities are also evident in cognitive capacities: see his disclaimers in *Psychosemantics* (1987: 152–3). But he takes the existence of any such systematicity in thought as the basic explanandum. Moving to a conclusion, he writes stirringly that 'the key to the nature of cognition is that mental processes preserve semantic properties of mental states; trains of thought, for example, are generally truth preserving, so if you start your thinking with true assumptions you will generally arrive at conclusions that are also true. The central problem about the cognitive mind is to understand how this is so' (1987: 154).

claim that thought is systematic is then said to be confirmed by the intrinsic connection between the ability to think one thought and the ability to think others. In Bechtel and Abrahamsen's example (1991: 212, modifying Fodor and Pylyshyn 1988: 39) 'anyone who can think *Joan loves the florist* can also think the *florist loves Joan*. For this to be so, "the two mental representations, like the two sentences, must be made of the same parts".'

But associationism and new connectionism fail to provide the requisite internal structure: therefore neither can be a genuinely cognitive theory. Associationist minds have no true intrinsic structure. The associationist strategy, in Fodor's picture, is to recognise in your ontology only 'the elements among which associations hold – ideas if you're Hume, stimuli and responses if you're Skinner' (1985a: 6–7):

if you say how the information in long-term storage is articulated, how the bits of information are connected to one another, then you provide the body of structural biases that somewhere has to be posited if one's going to explain the complexity of human achievement. That is good associationist doctrine. The structure that an associationist allows is the relational structure that holds among mental contents: not the structure of the mind, not the structure even of memory, but the structure of the things in memory. And since, of course, associationists are empiricists, they thought that [the] structure of the contents of mind reflects the structure of experience which in turn reflects the state of the organism's environment. That's what an associationist is: He's a guy who thinks that the apparent structure of the mind is really just the structure of the world. (1985a: 8)⁶

Because associationist organisms are slaves of context or creatures of the environment in this way, there is no *guarantee* that what is in their minds will develop the kind of constituent structure which, according to Fodor, is required to support cognitive capacities of inference, reason, and judgement. Learning, for an associationist, is statistical modelling sensitive to the frequency of, for example, co-occurring items presented in experience (Fodor and Pylyshyn 1988: 31). This leaves open the possibility that in a hostile environment, or in an unfavourable context of meagre stimuli, such minds would be unable to generate new combinations of existing items. This is absurd: a theory which allows even the possibility of 'punctate minds', minds which, for example, consist only of 'the ability to think seventy-four unrelated thoughts', without being able to relate them by way of their combinatorial constituent structure, has 'gone not just wrong, but *very profoundly wrong*' (1988: 40). It is not enough to say (Smolensky 1987/1991) that

6 Fodor quotes the learning theorist E.L. Thorndike's dissolution of memory to exemplify this exiguous explanatory strategy: 'there is no memory to hold in a uniformly tight and loose grip the experiences of the past. There are only the particular connections between particular mental events and others' (cited from Kline 1970: 662).

connectionist architecture is compatible with systematicity, that 'all biologically instantiated networks are, de facto, systematic' (Fodor and Pylyshyn 1988: 50; compare Fodor and McLaughlin 1990; and Clark 1993: 172). Since, on the connectionist 'no-structure' story, 'gaps are the unmarked case', connectionists require a separate explanation of the 'blatant' systematicity of 'inferential competencies': so there must be stipulated a mechanism which 'insures' this systematicity (Fodor and Pylyshyn 1988: 49–50). Only an architecture like that of RTM which postulates constituent structure as a primitive property of mental representations can thus *guarantee* systematicity. No diachronic account of contingent evolutionary and experiential history will count: it is just impossible 'to reconstruct the semantical coherence of thought' without postulating structure-sensitive operations on structured mental representations: this is 'the problem on which traditional Associationism foundered, and the prospects for solving it now strike us as not appreciably better than they were a couple of hundred years ago' (1988: 67).

There are two broad strategies of response to such criticism (see, for example, Chater and Oaksford 1990; Niklasson and van Gelder 1994). The associationist can accept the Fodorean requirements on cognitive theory, and try to mould connectionism so as to satisfy them: or the requirements themselves can be questioned. These are not exclusive strategies, and leaning different ways may be appropriate on different issues (chapter 17 below). Connectionist minds, for instance, are not such hostages to the environment as the traditional sensationalist interpretation of associationism would have it, and can introduce a 'body of structural biases' which is not dependent on present stimuli. Like the pores and spirits in Cartesian cyborgs, neural nets pick up and retain complex patterns over time, allowing considerable freedom from the moment.

There are empirical attempts to model constituent structure in connectionist representations; but other considerations can at the same time be brought against the central place given to systematicity and to the analogy with linguistic capacities. The non-existence of punctate minds can be ascribed to causes other than the mental structure in a language-like representational medium which RTM postulates. The priority of abstract reasoning as an explanandum is open to question. But there are other motivations for not playing Fodor's game. Mike Oaksford and Nick Chater critically examine the 'logicist' description (Fodor and Pylyshyn 1988: 29–30) of (classical) cognitive science as 'an extended attempt to apply the methods of proof theory to the modelling of thought'. Oaksford and Chater argue (1991: 2, 8) that this view of cognition as 'an implemented formal logic' is too rigid to account for looser cognitive processes which are 'inherently revisable', as any particular conclusion can change on the addition of new

information.⁷ I go on to identify further consequences of the logicist desire to model cognition on formal logic. An anti-logicist movement in cognitive science, drawing on distributed representation (Oaksford and Chater 1991: 33–4), will have to overthrow additional assumptions about cognitive control and executive mechanisms, and about the accuracy and adequacy of particular pictures of cognitive processes over time. It is important, then, to examine the fringes of Fodor's thought, by way of a slightly light-headed detour back through Coleridge.

11.2 Coleridge and Fodor on cognitive discipline

Samuel Taylor Coleridge had once been a Hartleian associationist and, going beyond Hartley, had believed in 'the corporeality of thought, namely, that it is motion' (Coleridge 1956–71: vol. 1, 137–8; Christensen 1981: 62–76). In *Biographia literaria* (1817), he attacked ugly consequences of Hartley's neurophilosophical associationism (BL VII: 218–23). The necessitarian laws of association would not only subordinate final to efficient causes in the human being, but would make the will and 'all acts of thought and attention' into mere 'parts and products of this blind mechanism' rather than what Coleridge claims they must be: 'distinct powers, the function of which it is to control, determine, and modify the phantasmal chaos of association' (VII: 218).

For Coleridge, the truth of associationism would be the greatest intellectual catastrophe in history.⁸ It would eliminate rationality, purpose, free will, the soul, consciousness, agency, choice and judgement, self, invention and creativity, art and beauty, prudence, ethics, and responsibility as well as theology. Although 'the excellent and pious Hartley' himself, after whom Coleridge had named a son, could surely not have been aware of 'the odium of these consequences', others 'in an unfortunate neighbour nation . . . have embraced this system with a full view of all its moral and religious consequences' (VII: 221). As antidote to evil French materialism, Coleridge quotes his own 1796 poem 'The

7 The specific kind of cognitive process which Oaksford and Chater suggest logicists cannot account for is 'non-monotonic' defeasible inference, in which 'it is possible to add premises and lose conclusions' (this contrasts with monotonic inference in which a conclusion, once derived, cannot be invalidated). Defeasible, non-monotonic inference, they argue, 'permeates every area of cognitive activity': if it 'cannot be elucidated within the logicist, proof-theoretic framework, then almost every interesting cognitive phenomenon will fall outside the scope of logicist psychological explanation' (1991: 2, 8–10).

8 This is how Fodor (1987: xii) would describe the collapse of common-sense intentional psychology in eliminative materialism. For Fodor, the end of the supernatural, a major component in Coleridge's fear, was much less disastrous than would be the collapse of belief/desire explanation. Coleridge believed that all the parts of the common-sense psychology of his day, including both the spiritual and the intentional, would stand or fall together.

Destiny of Nations', and vigorously describes how to treat an eliminativistic associationist. There are proud and mean people, the poem announces, who cheat themselves

With noisy emptiness of learned phrase,
 Their subtle fluids, impacts, essences,
 Self-working tools, uncaus'd effects, and all
 Those blind omniscients, those almighty slaves,
 Untenanting creation of its God!

Coleridge continues in prose: 'Such men need discipline, not argument; they must be made better men, before they can become wiser' (VII: 221).

Those who push too far their belief in the powers of subtle fluids or animal spirits 'that etch and re-etch engravings on the brain' (V: 211) are too far beyond the pale for debate, and must be forced into line. Their mistake, in Coleridge's view, is in not taking the controlling and determining powers of will and reason as given in psychological explanation. They give in to phantasmal chaos, to the despotism of the senses and the confusion of senseless memory, and they make all of us victims of jumbled association, in a state of 'complete light-headedness' (VI: 216).

Coleridge reports at length an 'authenticated case' in which will and reason are partly suspended, of the light-headedness to which associationism would reduce us all. An illiterate young woman in a Catholic town in Germany, 'seized with a nervous fever' or 'possessed by a very learned devil' (more likely, because of the 'known fact that she was or had been a heretic'), incessantly raved in Latin, Greek, and Hebrew (languages she had never learnt), in sentences 'coherent and intelligible each for itself, but with little or no connection with each other' (VI: 216–17). A 'young medical philosopher' traced her past to find that she once lived with an 'indulgent' Protestant pastor 'as his housekeeper': he read aloud 'out of his favourite books' while walking up and down outside the open kitchen door. Coleridge's diagnosis of the 'feverish state of the brain' in this case illustrates fear of the babble to which we would all be reduced if Hartleian distributed representation really was the ground of cognition. If 'every partial representation recalls the total representation of which it was a part', or 'any one part of [an] impression might recall any other part', the only practical result would be 'mere lawlessness' (VI: 215). Without the discipline imposed internally by the controlling will, only moral discipline will stave off jumble.⁹ Yet, as for Isaac Watts (chapter 9 above), surrender is intriguing too:

9 Christensen (1981: 110–12) questions Coleridge's aim in telling a story 'which, if it performs any function, reinforces the position of his adversary, albeit in a pathological extreme', and suggests that 'the maidservant in her feverish glossolalia' actually resembles Coleridge himself as writer, inspired by his reading into philosophic collage, a possessed ventriloquist overpowered by language.

in 1802, a trance-like vision while climbing had led Coleridge to speculate on such lightheaded states:

when the Reason & the Will are away, what remains to us but Darkness & Dimness & a bewildering Shame, and the Pain that is utterly Lord over us, or fantastic Pleasure, that draws the Soul along swimming through the air in many shapes . . . (in Holmes 1990: 330)

By 1817 such delightful torments were, officially, to be strictly warded off by asserting the primacy of self.

A perceived need for discipline in cognitive theory drives certain positions too in modern cognitive science, not as background assumption but at the centre of the issues. Again, I focus on Jerry Fodor's picture of cognition and human nature, set out by him with clarity, wit, and eloquence: and again, I am engaged here not in mounting a case against Fodor, but in tracing how his various commitments hang together.

Fodor's central contention was that 'what associationism missed . . . was the similarity between trains of thought and arguments' (1985b/1991: 39). He gives an example from Conan Doyle's story 'The Speckled Band' in which Sherlock Holmes describes to Watson the reasoning which led him to suspect that 'the Doctor did it with his pet snake'. Reminding us that philosophers are now allowed to treat reasons as causes (Davidson 1963; Fodor 1968a: 32–48), Fodor suggests that the reasoning described by Conan Doyle as characteristic of Holmes, giving premises which allow an inference to a conclusion, is a particularly fine piece of true reconstructive psychology.

Because this train of thoughts is tantamount to an argument, Holmes expects Watson to be convinced by the considerations which, when they occurred to him, caused Holmes's own conviction. (Compare the sort of mental history that goes 'Well, I went to bed and slept on it, and when I woke up in the morning found that the problem had solved itself' . . . That's mental causation perhaps; but it's not thinking.) (Fodor 1985b/1991: 40)

True thinking is like arguing: mere mental causation, the mere causal history of cognitive processes, is no use unless the thoughts which figure in the causal history leading to another thought also provide good grounds for the subsequent thought. This, for Fodor, is how we profit from thinking: true inferences generally suggest other true inferences (1985b/1991: 40–1). This is what the associationists missed, and so, Fodor judges, 'Conan Doyle was a far deeper psychologist – far closer to what is essential about the mental life – than, say James Joyce (or William James, for that matter)' (Fodor 1985b/1991: 41).

This is telling literary criticism. In *Psychosemantics* (1987: 18) Fodor expands. The normative requirements on cognitive theory entail a vehemently anti-associationist aesthetics:

Exactly what was wrong with Associationism, for example, was that there proved to be no way to get a rational mental life to emerge from the sorts of causal relations among thoughts that the 'laws of association' recognized. (See the concluding pages of Joyce's *Ulysses* for a – presumably inadvertent – parody of the contrary view.)

Molly Bloom's dispersed discourse at the end of *Ulysses*, mixing memory and desire, her remembering shot through with bodily intrusions (Joyce 1922/1968: 659–704), from Fodor's point of view merely parodies the essence of the mental life. This psychology could not be more different from Conan Doyle's. In Molly's night thoughts, any part of an impression can recall any other part, and so there seems (to a rationalist) no real difference between her train of ruminative cognitions and the 'mere lawlessness' which Coleridge saw in the German woman's raving.

Fodor's wish for cognitive discipline is at work again in his modularity thesis, which divides specialised, genetically determined, superficial, and relatively autonomous 'modules' from deep and rather mysterious central systems. These specialised 'vertical faculties', operating only in limited domains, are what cognitive science can potentially understand (1983: 38–46, 126–7), because their computations do not involve accessing all the information available to the organism. They are 'isolated' in being 'computationally encapsulated' (1985a: 12): so scientific psychological explanations can refer to modular processes without having to refer to everything else in the system (Fodor 1989/1990). What is good, for Fodor, about modular 'vertical faculties' is that 'these homunculi are disciplined. They're not allowed to merely babble. Gall's picture of the mind is not one in which everything connects, in which everybody shouts at the top of his voice, and the guy who produces the migraine first wins' (1985a: 12). Where Joycean mental causation is chaotic babble, to be avoided on pain of headache, disciplined modules provide a basis for rational Holmes-like thinking.

But disciplined homunculi are not enough. Fodor notes that Gall was attacked for dissolving the unity of the soul. There is a 'real issue' in the quest for something to integrate the information provided by all these modules which do not talk to each other, a problem to which Gall had no good answer (1985a: 14). But Fodor thinks that he has, and says that Gall's faculty psychology and his critics' belief in the unity of the soul 'were both right' (1983: 118). After God (evolution) made 'bundles of these fast superficial stimulus-driven mechanisms', says Fodor, he 'created the soul to put them together' (1985a: 18).

The central integrating system is organized in a way that is fundamentally different from the way that input systems are . . . The reason that the central integrator guy is there is the reason that the soul is always there; namely to

solve Gall's problem about what puts the peripheral stuff together. (1985a: 14–15; compare 1983: 101–19)¹⁰

It is no surprise, then, that these higher central processes turn out, on Fodor's view, to be all but inaccessible to science. They exhibit the creativity, non-encapsulation, holism, and 'passion for the analogical' which the high and puzzling processes of the soul were always meant to (1985c/1990: 202). Practically nothing, thinks Fodor, is known of what happens to information after the modules have got it 'into a form appropriate for central processing. . . . The ghost has been chased further back into the machine, but it has not been exorcised' (1983: 127).

Fodor, then, embraces a central executive with a vengeance. While he acknowledges that this may seem a gloomy as well as a tendentious evaluation of the state of cognitive science (1983: 127–9), the rigorous separation of disciplined homunculi from mysterious centre has, for him, distinct advantages. As Fodor regularly informs his audiences, the thing is that he hates relativism, and, sadly, he finds relativism 'affronting intellectual dignity' in far too many areas, permeating the views even of some otherwise respectable philosophers who, he complains 'appalled', tend to take holism, and thus varieties of relativism, for granted these days (1990: xii). His belief that various kinds of holism, both semantic and psychological, lead directly to relativism (1990: xi–xii) leads him to address the arguments which make holism attractive. Only atomist theories of meaning, in which items have their meaning in isolation and can retain the same meaning across uses and contexts, guarantee the possibility of genuine disagreement and democratic argument about the truth (Fodor 1995).

Quine, Goodman, Kuhn, Feyerabend, Putnam, Paul Churchland, 'and many others' make the hateful move from holism to 'anti-Realism (or relativism or Instrumentalism)' (Fodor and Lepore 1992: 11, 211). But Fodor's most extensive efforts in this direction combat the ideas that cognition saturates perception and that theory saturates observation. The modularity thesis separates the work of the vertical faculties, which include the perceptual modules, from the integration of perceptual information which, as Keith Campbell has expressed Fodor's view (1988: 168), 'is the work of the mind's Highest Command'.¹¹ The theory-neutrality of observation, which Fodor needs if he is to avoid thinking that scientific disputes 'are settled by appeals to coherence, or convention or – worse yet – by mere consensus', is retrievable if modules are insulated and iso-

10 Fodor has another argument for a central executive in cognition (1986: 320–1), where he acknowledges that associative networks do not need one. Cam (1988) challenges the idealisation of rationality on which these arguments rest.

11 Campbell's paper is a thoroughgoing application of Fodor's modularity thesis to 'a new and modest vindication of common sense' (1988: 161) against displacements of common sense's epistemic privilege by Quine, Feyerabend, and Paul Churchland.

lated from context and background knowledge (Fodor 1984/1990: 251; see also Fodor 1988/1990). Only thus can 'our agreement on the general character of the perceptual world . . . transcend the particularities of our training and go as deep as our common humanity. Granny and I hope this is so since common humanity is something that we favor.' At least some common-sense judgements, what Campbell dubs the 'Basic Observational Fragment of common sense', can be used regulatively, to impose 'definite, if mild, discipline' on the excesses of theory (Campbell 1988: 170, 174).¹² Within this fragment, common-sense judgements are the same for everybody, 'independent of theorizing . . . preference or prejudice, and not determined by social circumstances' (including, presumably, historical, racial, or cultural situation): if a person is 'deficient in judging of these sorts of things', then they 'are regarded as seriously deficient indeed' (Campbell 1988: 174, 169).¹³ Good liberals must uphold common humanity against dangerous jargon-benders who claim that mere persuasion and consensus, and thus not reason, drive conceptual change.

Fodor draws critical conclusions from his arguments against the saturation of perception by cognition. He notes conceptual and historical connections between this view and the corresponding determination not only of observation by theory, but also of values by culture, of science by class affiliations, and of metaphysics by language.¹⁴ All these views are relativistic in suggesting that 'rational criticism of scientific theories, ethical values, metaphysical world-views, or whatever can take place only within the framework of assumptions that – as a matter of geographical, historical, or sociological accident – the interlocutors happen to share. What you can't do is rationally criticise the framework' (1985c/1990: 205).¹⁵ Fodor hopes that the truth of the modularity thesis will at least be a step towards showing that relativism is false. He upholds

- 12 This is the fragment of common sense which contains basic categories 'marked by very high, unproblematic, intersubjective agreement' which 'enjoy plainly operable check-up procedures' (1988: 169). It provides a non-foundational yet epistemically privileged 'body of Appearances to which theories must defer and against which they may be tested' (1988: 172).
- 13 Campbell acknowledges that only empirical research can settle 'just which concepts prove to be basic categories for human beings', but gives some 'likely candidates – parts of the body, the terms by which familiar household artefacts and tools of trade are requested, bought and sold, the genres (roughly) of biology and geology, explorers' terms for landscape features' (1988: 169). Presumably it is only bad, relativistic theory that leads the anthropologists and psychologists who conduct such research to doubt the universality of classification schemes in these 'basic' categories: but, on the other hand, on the Campbell/Fodor view, their observations are not tainted by bad theory! Who is to go out and prove the existence of such unproblematic agreement?
- 14 The view Fodor opposes slips, in his characterisation, between the comprehensive *determination* of (e.g.) values by culture or observation by theory, and the weaker, more plausible *permeation* of values by culture or observation by theory, which would leave room for the involvement of other causal factors, whether internal or external to the agent.
- 15 This target is a very extreme incommensurability. There are many intermediate positions which take seriously the influences of theory on observation, class on science, and so on,

what relativists deny, the 'fixed structure of human nature', by insisting on 'the heterogeneity of cognitive mechanisms and on the rigidity of the cognitive architecture that effects their encapsulation. If there are faculties and modules, then not everything affects everything else; not everything is plastic' (1985c/1990: 205–6). The plasticity of mind is a myth: objective information is fed back by relatively stupid domain-specific faculties to an integrating central system which does the strange work of the soul (compare Fodor 1989/1990: 222–5).

Excessive dynamics in cognition would threaten realism, undermining the rationality on which social systems depend: so Fodor seeks cognitive theory which reinstates proper boundaries within the mind. Associationism is not only overly holistic, but also drops the mind's 'central integrator guy' who controls neatly bundled items passed on through the modules. The absence of control in associationism is a target for all its critics. Who will decode the babble?

without contending that criticism from outside the framework is impossible. Historians of science and ethnomethodologists distinguish the versions of a practice or set of beliefs given in 'members' accounts' and those given in 'strangers' accounts' (Shapin and Schaffer 1985: 4–7). If the extreme incommensurability thesis which Fodor attributes to his villains here were accepted, no historical or anthropological study of beliefs which we do not accept could ever get off the ground: there could never be an account by a stranger.