

The puzzle of elimination

... the Human Body is a Machin of an infinite Number and Variety of different Channels and Pipes, filled with various and different Liquors and Fluids, perpetually running, gliding, or creeping forward, or returning back . . .

(George Cheyne 1733/1976: 3–4)

Introduction

Remember Locke: the way a single human can, through continuous consciousness, incorporate past and present memories may be ‘an argument against those who would place thinking in a system of fleeting animal spirits’ (Essay II.27.13, my emphasis). It comes as no surprise to find the animal spirits which threaten such continuity disappearing as theoretical entities from roughly the 1740s onwards. The properties of spirits theory which were discarded in the conceptual change were, I suggested in chapter 8, as much metaphorical as physiological.

Making sense of the eventual rejection of spirits theory is a difficult task, which becomes visible only after dropping the assumption that animal spirits just obstructed neuroscientific progress. The spirits survived long into the Enlightenment, in the face of powerful objections from at least the 1660s onwards. What changed in the second quarter of the eighteenth century to make their elimination possible? Answers are hard to find, and mine are still badly formed. I consider five hypotheses: (1) that the elimination of animal spirits was inevitable; (2) that alternative theories displaced them; (3) that they were not eliminated but reduced (to nervous electricity); (4) that they were moralised out of existence as a function of the end of the dynamic aesthetic of the Baroque; and (5) that they became unable to support increasingly narrow conceptions of personal identity. Except for the first, these hypotheses are not mutually exclusive. This is all speculative, dealing as much with historiographical patterns as with historical truth, for there is a huge amount of source material still to survey.¹ I demonstrate the existence of the puzzle without conclusively solving it.

A couple of cautious quotations set the scene. Verbeke notes (1945: 506), at the end of a classic survey of ancient *pneuma* theories, that the doctrine of

¹ For orientation in primary texts see Rousseau 1969/1991, 1989; Clarke 1978; Yolton 1984b: ch. 8.

spirits, 'remaining vigorous until the seventeenth century, was gradually undermined by the experiments of Swammerdam, Galvani, and Volta' (compare Oppenheim 1991: 79). Vivian Nutton remarks (1981: 331) that the 'doctrine of animal spirits survived Vesalius' demonstration of the absence of the *rete mirabile* in man, and was not wholly replaced as a physiological explanation for sensation and action by Galvani's theory of animal electricity'. Verbeke and Nutton are more careful than those who assume that spirits were simply eliminated as neurophysiology at last became truly scientific. But their guarded statements reveal the magnitude of the historical problem. If animal spirits theory was 'not wholly replaced' by animal electricity, what did happen to it? If experimental evidence 'gradually undermined' nervous fluids and spirits, why did they succumb only when they did?

Hypothesis 1: inevitable elimination

In chapter 2, I quoted historians who wish the spirits had never retarded physiological theory: neuroscientists had to drop untestable hypotheses before their science could reach maturity. Brazier (1958: 198) complains that Galen's paralysing authority left spirits theory so sacrosanct that 'no one seems to have tested [it]' before the late seventeenth century: the picture of animal spirits and hollow nerves 'had to be dislodged' before electricity could be recognised as a factor in the transmission of nervous activity. But this last claim in fact points to a historical problem with the teleology which such language of inevitability implies.

The disappearance or fading of animal spirits from physiology occurred well before the development of a clear replacement. There was no consensual alternative for the radical young to rally round, pushing through a paradigm shift by dropping woolly old spirits. Electrical neurophysiology was little more than the dream of a theory at the time that theorists started to desert the spirits, as recent historians of animal electricity agree. The analogy between electricity and the nervous fluid was not a promising research subject before Galvani's experiments in the 1780s (Kipnis 1987: 111): Marcello Pera claims in *The Ambiguous Frog* (1992: 62) that the theory of neural electricity before Galvani was 'no more than a feebly substantiated conjecture'.

The disappearance of animal spirits was not inevitable: it was caused neither by the powerful objections to them, which had existed for many years, nor by the rapid development of a replacement, for the successor theory was uncertain.

Hypothesis 2: alternative theories

Even if need for a replacement theory was felt, most candidates were just as mysterious as spirits (Mazzolini 1991: 81–2). To say that animal spirits were replaced by nervous fluids (Clarke and Dewhurst 1973: 85) or abandoned in favour of nervous ether (Carlson and Simpson 1969: 108) is not enough without

specifying the relevant new theoretical framework more fully. This hypothesis is that animal spirits were superseded by theoretical entities deriving from newly attractive alternative physiologies, 'Newtonian' and 'vitalist'. But neither alternative was intrinsically more advanced, nor did they immediately open up new experimental domains.

Hartley and other 'Newtonian' physiologists shifted away from spirits towards ether and vibrations during this key period.² Cheyne retained animal spirits in 1715 but in 1740 rejected them and all kinds of subtle matter in favour of an elastic self-moving nervous mechanism described in terms of undulations and tremors (Yolton 1984b: 179). Browne Langrish, too, accepted animal spirits in 1733, while acknowledging that their unobservability raised doubts: but in 1747 he decided against nervous fluid, since ethereal nervous matter is moved by vibrations caused by an immaterial mind (Yolton 1984b: 174, 180). David Kinneir (1739: 11–19) explicitly set out to metaphorise usage of the term: any reference to a lively person's good spirits is really 'saying no more, than that such a one enjoys a good state of health' (1739: 11–12). But there was insufficient positive momentum to 'Newtonian' physiology, and its faltering progress in the 1730s (Brown 1968/1982: 308–53, 1987: 644–6; Suzuki 1992: 73–121) makes it hard to construe as a powerful research programme. Easy transfer from physics, not a coherent Newtonian alternative in physiology, partly motivated these changes of mind.

Another candidate alternative framework, which I can only mention, was the set of physiological beliefs which gained ground in a mid-century shift 'from mechanism to vitalism' (Brown 1974; compare Schofield 1970: 191–209; Roger 1980: 270–8; Shapin 1980: sect. 5; Wright 1985; Taylor 1988: 66, 80–1). Animal spirits were not necessarily incompatible with the reintroduction of soul into physiology, but they were often associated with Cartesian mechanism which denied the relevance of soul to life functions. Forms of 'vitalism' were motivated by desires to ward off the reductionist materialism implicit in Cartesianism, and to refute the atheistical identification of soul with animal spirits which materialists defended (Vartanian 1953: 216–41; Yolton 1991a). But there is no reason to believe that this forced the end of spirits theory, which had been accepted for centuries by many who were not materialists. The problematic capacities of the spirits, which seem sometimes to require intelligence of their own, were in no way intrinsically unusable by vitalists.

Hypothesis 3: reduction

One reason to query confidence in these explanations of the spirits' demise is that it is not clear that they really did disappear. Some historians of the spirits

² In the late 1740s, Hartley claimed to accommodate all hypotheses about animal spirits into his 'Newtonian Hypothesis of Vibrations' (OM 21, I.i.1, prop. 5; chapter 13 below).

fail to distinguish between two possibilities on the spectrum of theory-changes. Animal spirits might have been *eliminated*, as misplaced composites useless to a better successor theory (Cussins 1993), unrelated to the hypothetical constructs of that better theory. Alternatively, animal spirits might have been *reduced* (and not eliminated), in which case they would (have been decided to) be *identical* or closely related to the new theoretical entities. So when historians say loosely either that spirits just *are* nervous or neuro-electrical impulses (Jefferson 1949: 699; O'Neil 1974: 65; Monro 1975: 61; Schwartz 1978: 32; Reiss 1996: 598), or that they were *overtaken* and *replaced* by electrochemical impulses (Spillane 1981: 112; Flanagan 1991: 3), they should not be taken at their word. The genuine question is about where on the spectrum between reduction and elimination the theory-change over the mechanism of nervous transmission from animal spirits to electricity should be placed. It is not obvious that belief in the electrical nature of the nervous impulse really ended the cheeky spirits' theoretical career.

So Clarke and Jacyna's claim (1987: 160) that theories of hollow nerves and animal spirits were simply 'overthrown by the research on animal electricity first reported by Galvani in 1791' only reveals their desire to separate pre-scientific from progressive views. It is undermined by Galvani's own view, which they quote (1987: 167), that his research actually *explained* the nature of the animal spirits: other historians of animal electricity describe the process as reduction by identification, 'the recognition that the nature of those classical animal spirits was electrical' (Schiller 1982: 1). Though the exact date is disputed, the idea that animal spirits might be *identical* with neural electrical fluid was tentatively suggested by Italian physiologists in the years around 1750 (Pera 1992: 57–60). In 1760 Fontana argued that 'at the moment we cannot decide the identity of the electrical matter and animal spirits' (Pera 1992: 59–60). But after the electric-fish experiments of the 1770s, Galvani interpreted his own work with frogs as proving the existence of an electrical fluid inherent in and specific to animals (Walker 1937; Home 1970; Stevens 1973: ch. 2). Hoff argues (1936: 167) that Galvani simply accepted 'the current theory' of animal spirits, 'and substituted animal electricity for animal spirits', while Kipnis (1987: 141) comments that, to his contemporaries, Galvani 'proved what people had for long conjectured: the existence of the "nervous fluid"'.

Historical judgements about whether particular theoretical entities were eliminated or were reduced by identification are, however, rarely easy. Often the truth lies between, and in this case more detailed examination of the sources is required to understand why some mid- to late eighteenth-century physiologists rejected, and others retained, the concept of spirit (Jackson 1970; Yolton 1991a). But even if the question of whether animal spirits were replaced or simply adopted and adapted does turn out to be merely academic, it remains true that any residual spirits discourse did lose the metaphoric excess of the

spirits' heyday. Later eighteenth-century theorists tended to greater caution in neural ontology (Wightman 1958), in a general slippage of neuroscience away from metaphysics. The reference potential of the relevant terms contracted, as animal spirits' old associations with violence, transience, randomness, and excess became less common. One reason for preferring an answer closer to elimination than reduction is that spirits theory did include a considerable irreducible remainder, being in part a set of culturally permeated 'descriptions' of the cognitive phenomena it was meant to explain, which could not be incorporated or assimilated into more sober successor theories. My last two hypotheses about the theory-change are related statements of one perspective on these wider spirituous domains.

Hypothesis 4: the end of the Baroque

Edwin Clarke (1968: 139–41), discussing the modern belief that neuronal axons are hollow after all, ascribes earlier rejections of hollow nerves in favour of 'inflexibly static nervous tissues' to nineteenth-century methodological rigidity: the 'recent replacement' in the twentieth century of this notion 'by a more dynamic conception suggests a tie with the seventeenth century . . . [and] the Baroque demand for movement'. Later twentieth-century physiology and the old hollow-nerve theory 'have fundamental, dynamic features in common'. The neurobiologist Jean-Pierre Changeux (1985: 67–96) chose 'Animal Spirits' as title for the chapter of *Neuronal Man* which seeks to replace 'a static description' of cerebral wiring with 'knowledge of a different, dynamic order'. But mere invocations of the Baroque in body history do little explanatory work (Duden 1993: 85). It is easy enough to sketch a story of how increasing resistance to turbulence and motion across Enlightenment domains played some part in the taming of the 'Volatile Oeconomy of the Brain' (Mandeville 1711/1976: 131), and I do so here using Foucault's history of madness. But the following paragraphs are as much another statement of the historical problem as they are a solution to it.

The cultural disappointment of eighteenth-century physiology, argued G.S. Rousseau (1969/1991: 1–2, 13–17), was its failure to unite imagination with animal spirits. The thwarting of a promised 'organic marriage' of the spiritual with the material left at sea those who, like Kant and Coleridge, required a strong sense of unity from their theories of mind. The deification of imagination in the art of early Romanticism and idealism was achieved, in part, by removing it from natural philosophical contexts, rejecting the chaotic fluids wandering through brain folds (chapter 9 above). Animal spirits fell victim to this rejection of multiplicity and mixture.

In the central chapters of *Histoire de la folie* (1972: 181–361, abridged in 1965: 85–198), Foucault evoked the end of inner turbulence, tracing the disappearance in the eighteenth century of 'the image, with all its mechanical and metaphysical implications, of animal spirits in the channels of the nerves' (1965:

126). His analysis of the disappearance of certain 'explanatory myths' hints that the advancing 'morality of sensibility' could not deal with the 'maniacal spirits' and their 'pernicious mobility' (1965: 146, 129, 126). Cognition and physiology had coexisted in the heyday of the animal spirits in realms which were both organic and ethical. In relation to spirit, 'the distinction we make between the scientific and the moral was unnecessary, positively unwanted' (Davie 1963: 59). But in the gradual removal of concepts like sympathy and sensibility (chapter 9 above) from physiology to ethics, the cognitive realm became moralised and its norms narrowed and hardened, its invisible multiplicity of organic forces unified and restrained (Rousseau 1991, 1993; Stafford 1991: 417–36).

In Foucault's scheme, the dynamic of organic and moral penetration, with its indiscriminately permeable body, initially gave way to a 'physiology of corporeal continuity', exemplified by the concepts of sympathy and the complicity of internal parts, and finally to an ethic of nervous sensibility which treated hysteria, for example, as a psychological illness proceeding from a moral fault. English-language commentary on *Madness and Civilization* has focused on the later stages of this scheme, and the consequences for nineteenth- and twentieth-century moral psychiatry (Still and Velody 1992). But in the full French version, and in related early writings like *Mental Illness and Psychology* (1962/1976), Foucault constructs a remarkable positive picture of the earlier theoretical and phenomenological framework, of the jostling of spirits down unmarked inner paths, and of the 'original exuberance' of the early modern 'experience' of madness (Foucault 1962/1976: 65).

If there is naive nostalgia in Foucault's invocation of 'the grovel of flesh in the void' (1965: 279), or in his 'curiously innocent and untamed conception of madness' (Barham 1992: 47), it is not just in describing the 'easy wandering life' of late medieval and Renaissance madmen on their ships of fools (1965: 3–37, 1972: 13–55; Derrida 1978a; Megill 1992), but also in the alien world of 'classical madness' after the 'great confinement'. Careful, grey scholarship might find a residue of positive absences, an excess to the experience of madness in the classical age which is not inevitably subsumed in the marks of reason (Foucault 1972: 190–2, 223–5; Sutton 1994b). Awareness of the animal spirits' complex past encourages attention to Foucault's attempts to use old *theory* (rather than the voices of the mad, impossibly compromised by reason) to articulate something about the 'pre-modern' body, before it was burdened by consciousness with 'its complex coordinations, its intentional openings . . . the tension of the will that adapts and orders the compulsions' (1962/1976: 16–17). A whirl of warring elements, that body was 'a volume traversed by incessant motion whose disorder is . . . a lawless whirlwind in a chaotic space', subject to 'the incoercible agitation of desires'. This body vanished with the 'emergence of *Homo psychologicus*, possessor of internal truth, fleshless, ironical, and positive of all self-consciousness' and of the relation between himself and himself (1962/1976: 87).

Just as 'cures' initially premised on the old physiologies of fluids lost their sense when retained and moralised in the regimes of moral philanthropy (Foucault 1972: 316–61, trans. 1965: 159–98), so the task of the Enlightenment male philosopher concerned with the care of his self became the imposition of docility on his own body fluids: he must come to see cognitive order as natural, denying that memory was reconstructive and vulnerable to interference and confusion. As Foucault explicitly stated, the polymorphous positive experience of the classical age owed more to animal spirits theory than to metaphysical dualism or to 'Descartes' definition of substances'. Body and morality were not split in the scientific revolution, the self not till later pulled apart from its flesh and its fluids: when these schisms did occur, in the early nineteenth century, they did not result from 'a renewed loyalty to the *Meditations*' (1965: 182).

Hypothesis 5: 'moral Man'

This provocative story can be put together with my account in chapter 9 of the spirits and the self. Even if the self-conscious responsible agents inhabiting quietened static bodies were historically constructed (Richards 1992: 50–4, 242–6, 372–3), they cannot be voluntarily displaced with invocations of turbulence. We inherit the need, as well as the demand, to undertake that progressive integration of past and present, that assimilation of history into identity, which was encouraged by Lockean assumptions about memory and connecting consciousness, and enforced by institutions reliant on the responsible agents which such assumptions produce. It came to seem that a psychophysiology which allowed cognition to rest on volatile inner fluids was incompatible with certain strong assumptions about self and agency. Personal identity, with all its implications for morality and responsibility, requires a connected narrative, founded on memory: so memory must not depend on 'something organic which waxes and wanes and sometimes disappears altogether' (Diderot 1769/1964: 155–6). Distributed representations do not endure unaltered over time, and the instability of the irregular spirits which compose them jeopardised the self which memory was meant to support.

This is a loose claim, requiring evidence in specific contexts. But here is one manifestation: the Scots physician Robert Whytt, who emphatically rejected animal spirits and mocked appeals to them (1765: v–vi), saw no reason to use 'the irregular motion, increasing derivation, repercussion, confusion, or hurry, of the animal spirits, in accounting for the symptoms of nervous hypochondriac or hysteric disorders' (1765: 8, in Yolton 1984b: 187, n. 14).³ Whytt appeals instead to sympathy, peculiarly appealing in Scotland (Lawrence 1979; Wright

3 Yolton (1984b: 168) juxtaposes with surprise this 1765 rejection by Whytt of the spirits with Mandeville's 1711 use and description of them; but it is no shock that texts written over fifty years apart should reveal different attitudes to the same hypothetical constructs. I am trying to explain how such changes occurred over time.

1990). The delicate stability of the 'man of feeling' would be undermined by the fluid machinations of mischievous spirits in his internal passages. If the spirits threaten to enslave the soul, then, in time, it is the spirits which must go. In Scotland at least, medicine was meant 'to calm the mind' (Cunningham 1990): when it failed to do so, when confusion, hurry, or irregular motion seemed to result from particular theoretical commitments, those commitments were vulnerable.

The animal spirits had an intriguing afterlife in metaphor and in the language of character (see the references to Fielding, Austen, and Disraeli in OED, s.v. 'animal spirits'), but they were tamed and officially dead, leaving historians to commend the progress of truly scientific neurophysiology without them. Docility did not, however, come easy. Polarities of order and confusion are still invoked in cognitive science to support particular views on what it is about cognition and memory that needs explaining. The triumph of stasis was partial, and those who fervently desire cognitive discipline must still, like Isaac Watts in sickness, deny the tumult attributed to volatile brains by theorists unnaturally seduced by confusion. The need for a self freely to enforce order on memory and body did not disappear with the animal spirits' demise.