'The brain is a book which reads itself'.
Cultured brains and reductive materialism from Diderot to J.J.C. Smart.

Charles T. Wolfe

Centre for History of Science,
Department of Philosophy and Moral Sciences
Ghent University
Associate, Unit for History and Philosophy of Science
University of Sydney
charles.wolfe@ugent.be

In Helen Groth and Chris Danta, eds., Mindful Aesthetics: Literature and the Science of Mind
(London: Continuum, forthcoming)

1.Materialism, brains and discontinuity

Günther Mensching once observed that materialism is a “discontinuous tradition,”¹ which does not proceed by a transmission of doctrines, modified from generation to generation. Rather, each period founds a form of materialism on new bases, from theology onto natural history and emergent biology in the Enlightenment, biochemistry in the nineteenth century, physics in the mid-twentieth century and neuroscience ever since. Additionally, the ontological status of the brain is recognized in the eighteenth century as a key challenge in articulating a form of materialism. I shall focus on this problematic status of the brain as a starting-point for materialists, notably Denis Diderot in the eighteenth century and the twentieth-century proponents of the ‘identity theory’ of brain and mind (hereafter IT). Here, the discontinuity is not due to changing intellectual or socio-cultural contexts (“it is possible to attend to contexts and to brains at once”²) but rather to differing attitudes towards the brain, which themselves imply differing conceptions of science, materialism and culture.³

¹ Günther Mensching, “Le matérialisme, une tradition discontinue,” in Materia actuosa... Mélanges en l’honneur d’Olivier Bloch, eds. M. Benítez, A. McKenna et al. (Paris: Champion, 2000), 525, 513 (translations are mine unless otherwise indicated).
Materialism has no essence or suprahistorical definition, as it reflects particular intellectual constellations (thereby rendering any “history of materialism,” like Lange’s, untenable\(^4\)), but most generally, it is the view that everything that is real is material, or is the product of material processes. More precisely, materialism takes two distinct forms: a more ‘cosmological’ claim about the ultimate nature of the world, and a more specific claim about how what is mental is in fact cerebral: how mental processes are brain processes. Now, both of these seem to indicate a privileged relation between materialism and scientific inquiry – a privileged role for scientific inquiry.

An amusing, mostly unknown linguistic residue testifies to the close relation between materialism and the sciences. Prior to becoming a philosophical term in the later seventeenth century with More, Cudworth, and Leibniz, the word ‘materialist’ originally referred to pharmacists, who prepared the *materia medica*: the traveller Pierre Belon noted in 1553 that “Les drogueurs ou matérielistes qui vendent ordinairement les drogues par les villes de Turquie, sont pour la plupart hommes Juifs.”\(^5\) Indeed, from materialists as druggists to the ‘identity theorists’ of brain and mind in the 1960s (whose name came from psychophysics), and the subsequent *doxa* of materialism as the handmaiden of neuroscience, the question of which science is in play is always relevant.

In the twentieth century, the predominant science in the vision of a privileged relation between materialism and scientific inquiry was physics. Materialism became synonymous with ‘physicalism’: the entities that were considered to be real were those described in the physics of the time. But what about the status of brains, including within a physicalist scheme? An uneasy alliance then appears: is the materialist a brain theorist or a metaphysician bringing the rest of the world into line with physics? As David Lewis noted in his defense of the IT, “a confidence in the explanatory adequacy of physics is a vital part, but not the whole, of any full-blooded materialism. It is the empirical foundation on which materialism builds its superstructure of ontological and cosmological doctrines, among them the identity theory.”\(^6\) In fact, it has been observed that much of the identity theory – in its blend of logic, semantics and physicalism – conspicuously left out the messy details of neuroscience altogether.\(^7\)

My topic is not physicalism but rather the second species of materialism, claims about minds and brains, within which I distinguish more active, ‘plastic’ visions of the brain from more ‘passive’ visions, the former being open to an integration of brain and culture, science and our ‘fiction-making’ capacity. Denis Diderot was one of the first thinkers to notice that any self-respecting materialist had to address the question of what brains do, and how much of our

---


mental, affective, intellectual life is contained therein. This is not simply the prototypical materialist attribution of thought to the brain, which occurs fifty-sixty years earlier in John Toland (“Whatever be the Principle of Thinking in Animals, yet it cannot be perform’d but by the means of the Brain”; thought is a property of the brain⁸), and Anthony Collins (consciousness is “a real Quality, truly and properly inhering in the Subject itself, the Brain, as Modes of Motion do in some Bodies, and Roundness does in others”⁹). But neither Toland nor Collins feel compelled to provide more neuroanatomic detail or speculation; implicitly, this is not part of the responsibility of the philosopher as they see it.¹⁰

In contrast, what is unique in Diderot may not be flagrant, after a generation of models of ‘neural complexity’, ‘uncertain systems’ and ‘self-organization’ (some of which even credit Diderot, as Gerald Edelman does, for his ‘network’ concepts of brain and nervous system¹¹): the recognition that the brain presents a special explanatory and ontological challenge to the materialist, in a way that atoms, trees or polyps do not. This recognition leads Diderot to formulate a new kind of materialism, in which the nervous system in particular and network models in general, such as the harpsichord, play a constitutive role.¹²

After this the topic grew stale – more elegantly, “the nineteenth-century nervous system was less dynamic”¹³ – with repeated, knee-jerk reiterations of ‘psychophysical identity’ by ‘vulgar materialists’ such as Ludwig Büchner and Carl Vogt, who obsessively asserted that knowledge of the nervous system fully explains mental life, most famously in Vogt’s crude formulation, “thought is to the brain what bile is to the liver or urine to the kidneys.”¹⁴ Here, scientific practice is collapsed into materialism as an ideology. In the 1950s-1960s, a group of primarily Australian philosophers, U.T. Place, J.J.C. Smart, and D.M. Armstrong (with contributions from Herbert Feigl and David Lewis), took up brain-mind materialism afresh, under the name ‘identity theory’, in the sense that they were arguing for an identity between mental processes and cerebral processes; they actually waver in between being brain theorists (with surprisingly little invocation of neuroscientific evidence) and being metaphysicians bringing the rest of the world into line with physics.

---

¹⁰ A different case is David Hartley’s ‘vibratory’ theory of mind and brain, in which small vibrations are impressed in the solid filaments of the nerves by external objects; sensations are transmitted by aetherial vibration to brain particles. Different vibrations represent different primary sensations, or “simple ideas” in the brain, which can become complex ideas through associations with other chains of vibrations (Hartley, Observations on Man…, 2 vols. [London: Richardson, 1749], I, 13-16).
¹³ Sutton, Philosophy and memory traces, 26.
¹⁴ Carl Vogt, Physiologische Briefe (1847), 14th ed. (Gießen: Rickersche Buchhandlung, 1874), XIII, 323.
If we contrast Diderot’s materialism with that of the identity theorists, several notable features emerge, chiefly that Diderot allows for a much more culturally saturated brain, which he describes as a book—“except it is a book which reads itself”; and that he expressed his materialism in the form of an experimental philosophical novel, Le Rêve de D’Alembert (1769). Rather than just extracting abstract claims from his writings, I note that Diderot presents materialism (definitely not just the handmaiden of science then) in a literary or a para-literary work, which plays on experimentalism both in science and in style.\(^\text{15}\) I focus on two cases of materialism, both of which invoke the brain and the status of cerebral processes, although very differently: Diderot and the ‘identity theory of mind’. But some attention to history is called for.

2. Fluidity and spirits

The move from eighteenth-century ‘neurophilosophy’ to twentieth-century philosophy of mind may seem a large historical leap, but there is not much philosophical reflection on brains in the interim. What then of brains prior to Diderot? Doubtless the most important earlier episode in the history of brains is the neuroanatomical work of Thomas Willis, Professor of Natural Philosophy at Oxford and a founding member of the Royal Society.

Willis’s 1664 De cerebri anatome is a fascinating and influential work, articulating an experimental program for a bold, unconstrained neuroscience, filled with animal spirits and typologies of souls, yet based on comparative anatomy. Despite propelling the brain to the fore, however, Willis does not specifically worry about its particular status, or its cultural embeddedness. That is, unlike Toland and Collins’s metaphysical assertion of mind-brain identity or Locke’s “I shall not at present meddle with the Physical consideration of the Mind,”\(^\text{16}\) Willis opens up neuroanatomy as an unbounded playing field, without committing to any dangerous philosophical considerations on matter and mind. However, his emphasis on fluids, fermentation, the chemistry of life and the mobility of animal spirits cannot be understated in terms of a yet-unwritten history of brain plasticity. Plasticity notoriously is a ‘go-between’ concept, spanning experimental neuroscience and social, psychological, linguistic accounts of development, and sometimes invoked to stress the brain’s uniquely adaptive character.\(^\text{17}\)


\(^{17}\) Compare Peter Huttenlocher, Neural Plasticity: The Effects of Environment on the Development of the Cerebral Cortex (Cambridge, Mass.: Harvard University Press, 2002) to Bruce Wexler, Brain and culture: neurobiology, ideology, and social change (Cambridge, Mass.: MIT Press, 2006). Work from Atsushi Iriki’s lab shows that
A useful distinction, to which I return below, is between more ‘dynamic’ and more ‘static’ visions of the brain. Willis definitely belongs in the former category, for he envisions the brain as a plastic, self-transforming, self-organising entity, while opponents (but also later localizationist neuroscientists) see the brain as more of a passive mechanism – a lump of matter “of a clammy and unactive Nature and Substance; [which] seems as far as we can judge of it to be a mere passive Principle, as to the Acts of inward Sensation and Intellection,” in the words of one Boyle Lecturer attacking Willis, or a mere “Cake of Sewet or Bowl of Curds,” unfit to perform our cognitive operations, in Henry More’s words. Robert Boyle seems to be directly responding to More when he declares that “there must be in the brain . . . far more of mechanism than is obvious to a vulgar eye, or even to that of a dissector”; this “seemingly rude lump of soft matter” which looks almost like “so much custard” in fact has “strange things performed in it, . . . partly by the animal spirits it produces…” A vision of brain privileging transmission, dynamism, fluidity, and motion, or even just allowing for “more of mechanism in it,” is rather different from a “clammy nature,” a “cake of sewet” or a mere “rude lump of soft matter.” For after all, like Aphrodite, animal spirits would not be among those hypothetical entities which meekly hang around waiting to be observed. Making sense of the bits of the world which move might require a certain shimmying. Just as in late twentieth-century sciences ‘nothing is less static than the nervous system’, so the baroque internal edifices of the early modern neurophysiology of spirits were maintained only by motion.


21 Sutton, Philosophy and memory traces, 26.
appropriations of neuroscience. Subsequently, this over-determination of the plasticity of spirits leads to an emphasis on the cultural plasticity of neurology itself: “like the modern nervous systems it sought to describe, neurology proved itself to be particularly susceptible to absorbing the impressions, sensations and contaminations of the broader cultural discourses in which it was immersed.”

If we combine this gradual recognition of a cultural sedimentation of nervous systems with the fascination with nerves as explanatory of various disorders, we get what G.S. Rousseau called a new ‘neuromania’ in the eighteenth century: “in some qualified senses … much of the eighteenth-century Enlightenment was one magisterial footnote on nervous physiology, a remarkable attempt to secularize cognition and perception through the brain and its vassal nerves.” An intriguing case of this neuromania, and a central one here, is Diderot.

3. ‘The brain is a book which reads itself’: Diderot and plasticity

Just as the present essay is not an exercise in the socio-cultural history of brains, similarly, it does not focus on the cultural embeddedness of ‘neuromania’ (including spirits, brains, nerves, fibres), however fascinating that might be. I accept that brains are culturally sedimented; permeated in their material architecture by our culture, history and social organization, and that this sedimentation is reflected in cortical architecture, as first seen by the Soviet neuropsychologist Lev Vygotsky in the early twentieth century. Vygotsky emphasized the brain’s embeddedness in the social world, arguing that early social interaction leads to synaptic modifications. As his collaborator Luria put it, “Social history ties the knots that produce new correlations between certain zones of the cerebral cortex.” In today’s language, “Our minds and brains are (potentially) subject to constant change and alteration caused by our ordinary developmental engagement with cultural practices and the material world.” This is materialism sensu stricto, as it is a description of the properties of brains. Similarly, in the case of Diderot, we have someone who is not only fascinated with literary incarnations of the spirits (like Sterne’s Tristram Shandy) but is also engaged in natural-philosophical reflection on bodies.

---

organs, nerves and brains – one is tempted to say ‘scientific’ reflection but the very name for this part of Diderot’s œuvre, is a problematic issue.

Diderot’s reflection on the activity of brains and minds spans a variety of publications, some of which we would regard as novels, others as scientific commentary, and others as plain ‘philosophy’. The question of genres in his work is an old favourite, also as regards the relation between literature and science. Key texts here are his 1749 *Lettre sur les aveugles* and his 1769 ‘dialogue’ *Le Rêve de D’Alembert*, both of which seamlessly combine scientific speculation and literary experimentation. For present purposes, however, most relevant in addition to the *Rêve* is his *Éléments de physiologie*, which he worked on during the last two decades of his life, and did not publish.²⁷ I shall first discuss the *Rêve* before turning to the *Éléments*, which contains the key statement on the brain as a book.

*Le Rêve de D’Alembert* was one of Diderot’s personal favourites amongst his works, and remained unpublished during his lifetime; he gave one copy to Catherine the Great as a gift. It is divided into three dialogues, all of which feature real individuals as fictional characters. The first, between Diderot and D’Alembert, covers traditional philosophical issues such as self and world, matter and thought, and God. The second and main dialogue involves the somnolent D’Alembert, the doctor Bordeu, and Mlle de Lespinasse. It is based on the conceit of a delirious D’Alembert, raving in his sleep (in fact letting dream-associations do the argumentative work for him as he accepts what he denied in the first dialogue), so that his companion, Mlle de Lespinasse, grows alarmed and summons the doctor to whom she reports D’Alembert’s utterances. The rhetorical ingenuity is that D’Alembert has essentially uttered a series of rather concise materialist pronouncements, which Bordeu proceeds to explicate, carrying the reasoning much further, including in various science-fiction-like thought experiments. Mlle de Lespinasse is somewhat shocked at some of these, primarily on moral grounds, but gradually begins, in Socratic fashion, to ask increasingly pointed questions. The third dialogue is shorter again, and involves only Bordeu and Mlle de Lespinasse, reprising certain issues from the second dialogue, including biological and social aspects of monsters, and our material and sexual nature.

The *Rêve* articulates a network model of sensibility and a metaphysics of living, sensing matter to buttress it. Sensibility (*sensibilité*, a.k.a. ‘sensitivity’²⁸) is the property of the network, which is described through a series of metaphors – a bee-swarm, a harpsichord, a spider… and in the *Éléments*, a crawfish:

> the entire nervous system resides in the medullary substance of the brain, the cerebellum, the extended spinal cord, and the ramifications of this substance throughout various parts

---

²⁷ The title comes from Haller’s influential physiology textbook, *Elementa Physiologiae* (1757-1766).
of the body. It is a crawfish [spider], the nerves of which are the legs, and which is affected in various ways according to the leg.\footnote{Diderot, \textit{Éléments de physiologie}, in \textit{Œuvres complètes}, eds. H. Dieckmann, J. Proust, J. Varloot (Paris: Hermann, 1975-), vol. 17, 355.}

The bee-swarm is more a metaphor of organismic unity (our organs compose a whole organism like bees compose a swarm); the harpsichord is a step closer to our key metaphor of the book:

This organic faculty, by internally connecting the sounds within it, produces and preserves the melody therein. Suppose that the harpsichord has the power to feel and to remember, and tell me if it will not know and repeat of its own accord the airs that you have played on its keys. We are instruments endowed with sensibility and memory; our senses are so many keys that are struck by surrounding nature, and that often strike themselves…\footnote{Diderot, \textit{Rêve de D'Alembert}, in \textit{Œuvres complètes}, vol. 17, 102. On the role of analogies in Diderot, including ‘network analogies’ of the nervous system, see Anne Beate Maurseth’s \textit{L'Analogie et le probable: pensée et écriture chez Diderot} (Oxford: Voltaire Foundation, 2007).}

This emphasis on sensibility should also be understood as underscoring how organisms are not like the passive recording mechanisms which fascinated earlier natural philosophers: as Diderot exclaims, “What a difference there is, between a sensing, living watch and a golden, iron, silver or copper watch!”\footnote{Diderot, \textit{Éléments}, 335.} Notably, “our key characteristics lie in our brains, not in our external constitution (organisation)” so that “in order to explain the mechanism of memory we have to examine the soft substance of the brain.”\footnote{Diderot, \textit{Éléments}, 326, 470.} Memory is a product of our organisation (a term that here means our overall physiological configuration): “What is memory? . . . A certain [kind of] organisation which grows, weakens and sometimes is entirely lost,” “a corporeal quality,” an “organic faculty,” the aggregate of all the sensations I have experienced.\footnote{Rêve de D'Alembert, 101; Éléments, 335.} This makes for a different kind of materialism, but so far, it does not seem that unusual: we are different from other material arrangements of particles because our key characteristics reside in our brains, which themselves are the locus of cognitive processes. What \textit{is} different is Diderot’s central choice of metaphor for the brain, in the \textit{Éléments}:

In order to explain the mechanism of memory we have to treat the soft substance of the brain as a mass of sensitive and living wax, which can take on all sorts of shapes, losing none of those it received, and ceaselessly receiving new ones which it retains. There is the book. But where is the reader? The reader is the book itself. For it is a sensing, living, speaking book, which communicates by means of sounds and gestures the order of its sensations; and how does it read itself? By sensing what it is, and displaying it by means of sounds.\footnote{Éléments, 470; cf. 237-238.}
Diderot may be influenced here by the mysterious, heterodox Benedictine monk Léger-Marie Deschamps – they met several times in 1769, when Diderot was writing the Rêve, and Deschamps shared with Diderot his ‘clandestine’ materialist treatise, La Vérité ou le vrai Système.\(^{35}\) Deschamps wrote, “to read me, to hear me read is to become composed of my work, which then acts physically by the eyes or ears on the fibres of the brain, and raises them to a given tone, according to its impressions on them.”\(^{36}\)

In fact, Diderot seems to move between different positions concerning the brain, at different times: either our key characteristics are located in our brain, or the brain is just “an organ like any other,” “a secondary organ” (467), “merely a secretory organ” (353); this wavering on whether the brain is ‘special’ or ‘an organ like any other’ occurs within the Éléments itself. It is likely that these shifting attitudes toward the ontological status of the brain map onto Diderot’s shifting views on whether or not one can defend any degree of ‘anthropocentrism’ within a naturalistic universe: that is, if the brain is special, it allows for a form of naturalism in which humans are unique, and if it is an organ like any other, we find ourselves in a more thoroughgoing naturalism – albeit one in which all of matter is, actually or potentially, living matter.

Similarly, Diderot sees that the concept of sensibility allows him to integrate the reactive, representational capacity of mind (the nervous system, the brain as a “book which reads itself”), while maintaining a thoroughgoing naturalism: there are no properties which are not properties of natural beings subject to causal processes as specified in the natural sciences, although these properties may vary: naturalism as physicalism is rather different from Diderot’s reduction to matter as the bearer of vital properties. Yet these ‘network’ properties are not properties of matter as such. Rather, they belong to a body-brain network (like the spider and spiderweb, which Diderot views as forming one organic system), which he describes as “a system of actions and reactions.”\(^{37}\)

But Diderot’s claims are not just ‘scientific’, even if there are naturalistic elements in both his experimental prose and his physiological ‘notebooks’. One approach views Diderot as a proto-Bachelardian poet-metaphysician of the cosmos\(^{38}\) (as in the Rêve’s “human polyps on Jupiter or Saturn”\(^{39}\)), who leaps into associative freedom, beyond the constrained empirical studies of scientists like Haller and others. Some describe this speculative dimension, in which Diderot’s scientific imagination can reach conceptual ‘places’ that science cannot, as a kind of science-fiction, or better, as Anne Vila does, as “a thought experiment on sensibility” – a thought

\(^{35}\) For Deschamps’ influence see Vernière’s note (Diderot, Œuvres philosophiques, ed. Paul Vernière [Paris: Garnier, 1961], 300, n. 1), and André Robinet, Dom Deschamps (Paris: Seghers, 1974), 45ff.


\(^{37}\) Diderot, Éléments, 337.


\(^{39}\) Diderot, Rêve, 125.
experiment which instantly has material effects and is itself “materialized.”

The other view, which focuses on the *Éléments*, is to view Diderot as a commentator on scientific studies of sensibility who remains at the level of fragments, unable to provide his own scientific theory – at best, a (materialist) philosopher accumulating information to support his metaphysics. I prefer the more sympathetic and expansive version of this view, which does not strictly demarcate his reflections on sensibility, brains, bodies and networks from a genuine, if speculative (vitalist) life science.

Read from this perspective, two key features of Diderot’s cultured-brain materialism appear more sharply.

First, that the inextricable relation between thought experiments and the ‘materialisation’ of such acts of the imagination is constitutive of his thought: he speaks of “undertak[ing] by means of thought what Nature does sometimes.” In an extended materialist universe (albeit one in which sensibility is a basic property of matter), imagination and what Diderot calls “supposition” play a key role: the *extension of materialist ontology*. Similarly, when he engages with debates on biological generation: Diderot neither dutifully takes notes on experimental scientific work, nor wildly speculates on it; he seeks to extend the reach of materialist ontology – here, as regards Life. Second, that he would not countenance a straightforward linear explanatory scheme in which knowledge of brains is more primary and more exact than knowledge of culture. He would not think that a future neuroscience would enable us to ‘explain’ cultural forms, as in some current ‘neuroaesthetics’, which tells us that in reading prose, “the line ‘He had leathery hands’ has just stimulated your sensory cortex in a way ‘he had rough hands’ can never hope to.’ His way of integrating ‘cultured brain plasticity’ (“the brain is a book which reads itself”) with broader imaginative extensions (“human polyps on Jupiter”) is not like neuroaesthetics, either in its positivistic incarnation (neuroscience *explains* art) or in its utopian incarnation (creating new circuits in art means creating new circuits in the brain, in Deleuze-Guattari’s phrase). It is less unidirectional and mechanistic than the former, more flexible in its explanatory schemes; but it is also less utopian than the latter. In this it partly resembles the ‘libertine’ and ‘clandestine’ appropriation of scientific discourse as found in the so-called Radical Enlightenment.

---


41 *Rêve*, 149.


44 The way in which eighteenth-century libertine literature, including Sade and Boyer d’Argens, appropriates ‘scientific’ discourse and bends it to suit its own mutated-materialist ends (vital fluids, animal electricity, humoral determinism…) has been studied, and is not the topic of the present essay. See e.g. Caroline Warman’s *Sade: from materialism to pornography* (Oxford: Voltaire Foundation, 2002) and the current research of Claudia Manta.
It is both striking and puzzling that the kind of dynamism on display both in early modern discourses of animal spirits and in Diderot’s vision of a self-transforming brain, is absent from the birth of experimental neuroscience in the nineteenth century. As Sutton politely notes, the nervous system here was less dynamic. But it is significant here, for it is a further attempt – and a novel one – to correlate cognitive function with bundles of “spatially structured living matter,” under the heading of localization. Most visible in this project is the ‘pseudo-science’ of phrenology, along with various assertions of psycho-physical or psycho-physiological identity – which belong more to a militant rhetoric of science than to the messy reality of empirical research. It is a very different kind of identity, or at least a different kind of identity claim, than that of the Australian materialists discussed below.

4. From the material organ of the mind to blotches on the wall: identity quarrels

The localization of mental functions in parts of the brain has been described as the closest meeting-point of experimental natural science and core issues of human nature, including the mind-body problem. More precisely, by the nineteenth century “most physiologists, physicians, and naturalists would [largely] agree that, whatever the function embodied and instantiated by some organ, the activity of this bodily part resided … in the portion of living matter by which it was subserved,” as in Ewald Hering’s 1870 address, “Memory as a general function of organised matter.” Yet the one ‘site’ where the correlation of bundles of matter and functional properties was not clear was the brain.

Franz Joseph Gall (1739-1828) sought to correlate detailed anatomical descriptions of the brain with morphological features of the skull, and with observations on the physical, measurable, nature of our mental and moral faculties – what he called ‘organology’ and his colleague Johann Spurzheim more enduringly termed ‘phrenology’ (a term Gall rejected). Each mental faculty possessed its ‘organ’, the development of which could be analyzed by examining the external features of the skull. For Gall, the brain was the material organ of the mind, composed of a definite number of compartments, each of which is the seat of an intellectual faculty or sentiment. Each mental function corresponded to a part of the cortex: the features of

---

45 Sutton, Philosophy and memory traces, 26.
46 Métraux, 164.
48 Métraux, 164.
the mind were determined by the configuration of the brain. The mind is seen as “a set of functions instantiated by spatially circumscribed and mutually connected portions of neural matter.”

Despite its somewhat nefarious ideological dimensions, phrenology and related projects are important because they explicitly defend an empirical form of mind-brain materialism, which naturalizes or at least partly demystifies ‘the mental’. These are reflections on the material basis of cognition, based on the experimental investigations of brains – portions of neural matter. Notice that this approach is light-years removed from Diderot’s ‘cultured-brain materialism’, which emphasizes plasticity and the socio-cultural embeddedness of the brain; but, like his approach and unlike the IT, to which I now turn, it is a straightforward set of empirical claims.

The identity theory (IT), the project to philosophically articulate the identity between cerebral processes and mental processes, is the paradigmatic expression of materialism in twentieth-century English-speaking philosophy. Here, the goal is to resolve or eliminate some of the problems that have arisen for philosophy as a result of the ‘mind-body problem’, in at least two disparate ways, which hang together, sometimes well, sometimes awkwardly: empirically (seeking to apply some of the scientific successes of previous generations to the status of the mind) and logically (what does it mean to say that a mental event might be the same as a brain event?). Crucially, the status of the brain here differs from Diderot’s “book which reads itself,” including with regard to the status of science.

Smart states a mild version of the ‘scientific’ approach (as we shall see, Place has a stronger version, bordering on scientism):

That everything should be explicable in terms of physics (together of course with descriptions of the ways in which the parts are put together – roughly, biology is to physics as radio-engineering is to electro-magnetism) except the occurrence of sensations seems to me to be frankly unbelievable.

Indeed, the question of what science the identity theorists found their claims on will be significant: is it neurology? molecular biology? physics (which yields the idea of causal closure)? This is distinct from the logical aspect of the question: when we say a sensation is a

52 Métraux, 183.
54 Smart, “Sensations and brain processes,” 142, emphasis mine.
brain process, are there two kinds of things which are related (and what is the nature of the relation), or is there only one thing? Of course, these can be unified, as in Armstrong’s presentation of the IT as a solution to the mind-body problem, which can account for (1) the unity of mind and body, (2) the numerical difference between minds, (3) the interaction of mind and body, (4) the emergence of mind (e.g. in evolutionary terms), and (5) inner states.\textsuperscript{55} But in fact the IT (as distinct from later neurophilosophy, which focuses more on points 4 and 5) chiefly works on the first three points, in very internalist terms.

What exactly is happening when I say ‘I see a orange blotch’? Smart’s 1959 article begins abruptly: “Suppose that I report that I have at this moment a roundish, blurry-edged after-image which is yellowish towards its edge and orange towards its center. What is it that I am reporting?”\textsuperscript{56} He later explains that he wants to show “that there are no philosophical arguments which compel us to be dualists” (143): if recent science shows that organisms can be understood as physico-chemical mechanisms, what about the mind? Smart suggests, cautiously, that we should start by rethinking the language we use to describe mental events, such as ‘I am in pain’ or ‘I see an orange blotch on the wall’. Our sensation and the corresponding brain-process might well refer to the same thing (151), even if the logic of the two statements may be different.

The IT seeks to strip the ‘mental’ dimension from my report of, say, a sensation of orange, so that it becomes ‘There is something going on within me like what happens when a certain physical stimulus is present’. However, Smart cannot commit to which physical event the mental event is like; it is a contingent identity. That is why the IT is topic-neutral, more of a place-holder than a robust materialist account, although he has left room for science to ‘fill in the gaps’. Indeed, as “brain-imaging evidence begins to replace the subject’s introspective report in determining the occurrence and nature of her conscious experience,” Place feels that the IT moves closer to being a genuine, verifiable scientific theory, just as for Feigl, a “detailed account of brain-mind identities is a matter of future progress of psychophysiological research.”\textsuperscript{57}

To Smart, this is partly right and partly wrong. He agrees that the claim ‘sensations are brain processes’ can be tested scientifically (recall Vogt’s ‘The brain secretes thought like the liver secretes bile’), yet he does not want to make empirical claims, not least because of the difficulty of correlating complex, intentional mental states with patterns of neuronal activity: “conscious experiences must be processes involving millions of neurons, and so their important likenesses and unlikenesses to one another may well be statistical in nature.”\textsuperscript{58} And statistical

\textsuperscript{55} Armstrong, 75-76.
\textsuperscript{56} Smart, “Sensations and brain processes,” 141.
\textsuperscript{57} U.T. Place, “We needed the Analytic-Synthetic Distinction to formulate Mind-Brain Identity then: we still do,” Symposium on ‘40 Years of Australian Materialism’, ms., Dept. of Philosophy, University of Leeds, 1997, 15 (thanks to Dr. M.-C. Wright of the University of Leeds for providing me with a copy of Place’s paper); Feigl, 90.
\textsuperscript{58} J.J.C. Smart, “Materialism,” \textit{Journal of Philosophy} 60 (1963), 656. Smart’s version of the IT is thus more flexible than Place’s.
correlation is an unlikely basis for identity. In the end, statistical correlations and contingent relations do not sound too boldly materialistic, unlike Diderot’s ontologically assertive materialism: he may have thought brains were special, _qua_ ‘books which read themselves’, but Diderot was nevertheless willing to make empirical, _material_ claims about living matter as a whole, including sensibility, cognition and mental life.

There is a surprising absence of scientific detail in the IT – of embodiment. There are hints of DNA and neurology; but mostly, there is physics: materialism post-1950s is _physicalism_. How does one get from physics to the specific relation between brain states and mental states? Armstrong says it is a “good bet”; in fact it is partly an inductive argument based on the past successes of reductionist approaches. As Bickle _et al._ note, there were various plausible candidates for psychoneural identities in scientific literature before the IT: Hebb’s _The Organization of Behavior_ (1949) explains psychological phenomena in terms of known neural mechanisms and processes. But the IT primarily focuses on the _logic_ of the ‘identity’ between sensations and cerebral processes (although as noted, Place differs from Smart on this). The contrast with Diderot and other materialists should be apparent: here, there is no strong claim that locates mental life in a bundle of (neural, material, sensing) matter. However much Diderot employed metaphors, he was speaking about the mode of existence of _real brains_. To be clear, the issue is not _how much_ is known about the brain: writing after Gall and Flourens, Place and Smart know, or could refer to, far vaster swathes of neurocircuitry than Diderot or _a fortiori_ Collins; but they chose not to.

Nevertheless, whether in the form of contingent identity claims (Smart) or of explicit empirical claims to be filled in by future neurophysiology (Place), the IT is, broadly speaking, _scientism_, i.e. a vision in which materialism is ultimately fatal to philosophy by serving as the handmaiden of science:

> [I]t would seem that the long reign of the philosopher as the professional in charge of the mind-body problem is finally coming to its end. Just as … the origins of the universe used to be a theological problem and is now an astronomical one, so the mind-body problem is about to pass from the grasp of the philosopher into that of the neuropsychologist.

---

59 Timo Kaitaro, “Brain-Mind Identities in dualism and materialism: a historical perspective,” _Studies in History and Philosophy of Biology and Biomedical Sciences_ 35 (2004), 629, 642. In his intriguing paper, which is orthogonal to this one, Kaitaro suggests that the IT matches Charles Bonnet’s dualism rather than Diderot’s materialism. He notes that dualists such as Bonnet, in addition to referring to the seat of the soul, were often ultralocalizationists regarding the anatomical correlates of separate ideas, for they considered that there was a specific fibre or fibres in the brain for each idea. However, “the postulation of such identities in itself is not committed to dualism or materialism” (Kaitaro, 629).

60 Armstrong, 90.

61 Place, “We needed the Analytic-Synthetic Distinction,” 16.
It should be clear that Diderot’s materialism is not of this sort, notably because of its speculative, fictional component. Indeed, materialism does not have to be a ‘handmaiden’ of the natural sciences; conversely, science is not necessarily “the laboratory of materialism.”

5. The brain between scientism and fiction

Materialisms diverge as they address the status of the brain, between more plastic and aesthetically open models (Diderot and others), and more static, quantitative models (including phrenology and the IT). The difference is not between a specifically cerebral materialism and more disembodied models; Toland and Collins also explicitly identify cerebral processes and mental processes, as does nineteenth-century localizationism in more detail (although it is not necessarily philosophically materialist) and psychophysics. Diderot is distinctive in recognizing a kind of cerebral plasticity and self-organization, and insisting on their irreducibly fictional dimension. Is he then the ‘predecessor’ of a more holistic trend running counter to ‘Australian’ materialism? This was claimed in once-fashionable works seeking to unify aesthetic complexity and physics, like Prigogine and Stengers. But if we make Diderot into the poster child of an ‘enchanted materialism’ of fiction and embodiment, we lose sight of his own reductionism. However, Diderot (like Vygotsky for whom “History, changing the human type, depends on the cortex; the new socialist man will be created through the cortex”) and some contemporary cultured-brain theorists, provides not just a materialist outlook, but one which acknowledges the self-organising dimension of brains.

One might say that Diderot’s is a materialism without physicalism, whereas the IT is a materialism without embodiment. Somehow, for reasons that are partly intuitive and affective (embodiment implies affectivity), it is embodied materialism that recognizes the ‘plasticity’ of the book which reads itself, in contrast to the older view of the brain as merely passive. Unlike the “clammy and unactive Substance,” the “meer passive Principle,” Diderot’s more plastic vision of the brain is neither passive nor mechanistic. Cultured-brain materialism is thus not a ‘scientism’ or a denial of the symbolic and valuative dimensions of life, contrary to the criticisms launched recently by David Hawkes – which are almost identical to the charges brought in a

---

63 Ilya Prigogine & Isabelle Stengers, Order Out of Chaos (London: Heinemann, 1984), 82-83.
66 Hancock, 243.
1933 article by Raymond Ruyer, tellingly entitled “What is living and what is dead in materialism.”

Ruyer suggests a thought-experiment: imagine a law court as seen through the eyes of a materialist. “The halo of meanings, essences and values” vanishes, and what is left is the “functioning of a sort of complicated mechanics” whereby brains produce articulations, which in turn generate vibrations in the air, and thereby modify other nervous systems. For Ruyer materialism is a strange kind of reductionism which denies the reality of social institutions, values, and of course minds. Hawkes’ similar vision has been nicely challenged by John Sutton and Lynn Tribble. They note (without any particular focus on the status of brains) that materialism need not claim that ‘only matter exists’, but is instead “firmly pluralist” in its ontologies: “Even if all the things that exist supervene on or are realized in matter, the materialist can still ascribe full-blown reality to tables and trees and tendons and toenails and tangos and tendencies”; an account including the brain need not exclude “memories, affects, beliefs, imaginings, dreams, decisions, and the whole array of psychological phenomena of interest to literary, cultural, and historical theorists.”

The materialism of the ‘cultured brain’ is very much of this sort: it integrates the brain and the affects, cerebral architecture and our aptitude to produce fictions. But it is not enough to rebut these ‘antelapsarian’ visions of a cold, dead materialism seizing living value, sentiment and meaning in its embrace and reducing them to piles of “clammy” matter (sewet, curds, etc.). For from the opposite side of the spectrum, some ‘neuroaesthetic’ thinkers claim to integrate materialism, brain science and art, but in the flattest way: “I picture a future for writing that dispenses with mystery wherever it can, that embraces the astounding strides in thought-organ research. Ideally, a future where neuroimaging both miniaturises and becomes widespread, augmenting the craft of authors, critics, agents and publishing houses.”

Despite its potential scientism, the IT is not as caricaturally restrictive as this; it can indeed consider that our knowledge of mental life is filled in by the progress of neuroscience (Bickle), or instead become a “heuristic identity theory.” Either of these can preserve narrativity as a supervenient feature, in which selfhood emerges out of stories my brain ‘tells itself’: a “narrative self” in which “consciousness is a property I have by virtue of my brain’s

---

68 Ruyer, 28.
70 Walter, “What neuroscience tells us about the art of fiction.”
attributing it to me. My story doesn’t have to cohere completely to be useful.” Of course, this narrativity is a far cry from Diderot’s way of preserving a role for fiction within materialism.

How is this role preserved? The embodied character of Diderot’s materialism means that the brain has particular properties distinct from physical nature as a whole. It is a brain embedded in culture, expressed in works where experimentalism is both ‘science’, literature and metaphysics. Does the cultured-brain materialist have to grant special ontological status to the brain? Clearly, for the IT, the brain does not have an ontology; there is physics, and anything above, including biology and neuroscience, is like special kinds of radio engineering. In contrast, for Diderot, it does. But how can materialism maintain that the brain has an ontology without reintroducing ‘kingdoms within kingdoms’ (as Spinoza described the erroneous belief that human nature is apart from the laws of nature as a whole)? If we over-ontologise the brain in order to not be mystical dualists or knee-jerk anti-scientists, we may also run the risk of reconfiguring humanity as just “a cerebral crystallization.”

References


74 Deleuze and Guattari, 197-198.


Sutton, John. and Tribble, Evelyn. “Materialists are not merchants of vanishing. Commentary on David Hawkes, ‘Against Materialism in Literary Theory’.” Early Modern Culture 9, forthcoming


