Materialism and “the soft substance of the brain”: the case of Diderot

Charles T. Wolfe
Department of Philosophy and Moral Sciences
Sarton Centre for History of Science
Ghent University

Longer draft (2015) of paper forthcoming in British Journal of History of Philosophy (a number of additional citations appear in published version)

Abstract

Materialism is the view that everything that is real, is material or is the product of material processes. It tends to take either a ‘cosmological’ form, as a claim about the ultimate nature of the world, or a more specific ‘psychological’ form, detailing how mental processes are brain processes. I focus on the second, psychological or cerebral form of materialism. In the mid-to-late eighteenth century, the French materialist philosopher Denis Diderot (1713-1784) was one of the first to notice that any self-respecting materialist had to address the question of the status and functional role of the brain, and its relation to our mental, affective, intellectual life. After this the topic grew stale, with knee-jerk reiterations of ‘psychophysical identity’ in the nineteenth-century, and equally rigid assertions of anti-materialism. In 1960s philosophy of mind, brain-mind materialism reemerged as ‘identity theory’, focusing on the identity between mental processes and cerebral processes. In contrast, Diderot’s cerebral materialism allows for a more culturally sedimented sense of the brain, which he describes in his late Elements of Physiology as a “book – except it is a book which reads itself”. Diderot thus provides a lesson for materialism as it reflects on the status of the brain, science and culture.

“Contexts without brains are empty, brains without contexts are blind”
(Present 2014)

For John Sutton

Introduction: typologies of materialism

The scholar of early modern materialism wishing to create some ‘reasoned’ typology (in the sense of an histoire raisonnée) out of the bewildering multiplicity of materialist sources, contexts, theoretical and metaphysical commitments, may consider Friedrich Lange’s nineteenth-century History of Materialism1 longingly, with its broad, synoptic perspective – leaving aside the fact that Lange ultimately wrote in order to refute materialism. Yet the possibility of such a synoptic view

---

1 Lange 1892 (original publication 1866).
must be tempered with the realization that materialism is, as Günther Mensching put it, a "discontinuous tradition": that is, there is no monolithic, progressively articulated position called ‘materialism’ comprised of an interconnected set of doctrines transmitted and modified from generation to generation. Rather, each period founds a form of materialism on new bases, e.g., theology in the sixteenth-seventeenth centuries, natural history and emergent biology in the eighteenth century, as I have discussed elsewhere; biochemistry in the nineteenth century, physics in the first half of the twentieth century and neuroscience ever since. Contrary, then, to nineteenth-century attempts such as Lange’s, there can be no separate history of materialism. Yet this should not lead us to jettison conceptual clarity in the face of historical and scientific diversity. For the sake of such clarity, I suggest that we distinguish, as some early commentators but also anti-materialist authors did, between two core forms of materialism, which I’ll term materialism$_1$ and materialism$_2$: a ‘cosmological’ materialism, which defends the thesis of the materiality of the world, in different versions according to different matter theories, and a ‘psychological’ or ‘cerebral’ materialism, according to which the mental is really the cerebral. The present analysis focuses on varieties of the latter, materialism$_2$. But a second distinction, applicable to both, is of equal or perhaps greater importance here: both of these forms of materialism allow of more passive (or static) and more active (or dynamic) variants. More passive conceptions of matter and materialism are often articulated by opponents wishing to emphasize the dangers in a reduction of (variously) mind, emotions, free will, morals and so on, to a ‘mechanistic’, passive universe of ‘dead matter’.

Debates over the reduction of various entities to arrangements of matter are an obvious feature in materialism$_1$, with a marked contrast between more mechanistic and more dynamic or vital matter theories. In more mechanistic versions of materialism$_1$, say, Hobbes, what is real is physically specified matter and motion, and the only kind of substance that exists is body thus specified, to which faculties such as memory reduce (memory will become relevant in later sections of this paper). A particularly crisp mechanistic statement is J.G. Walch’s entry on ‘Materialism’ in the 1726 Philosophisches Lexicon: in materialism, “all the occurrences and operations of natural bodies are derived from the bare properties of matter, as from its dimension, shape, weight, confrontation and mixing, and thus will not admit any other spiritual principle except for the soul’s”; “but,” Walch adds, “that is exactly what is called Mechanism.” The ultimately mechanistic passivity of matter was also an anti-materialist commonplace: for Pierre Nicole (one among many), “matter lacks any internal cause of its existence . . . it is ridiculous to attribute to the most vile and desppicable of all beings, the greatest of perfections, which is to exist by oneself” (Nicole 1671 in Nicole 1714, 27). Matter is necessarily passive and hence cannot explain forms of ‘activity’ such as mind and by extension free action.

---

2 Mensching 2000, 525, 513.
3 Wolfe 2014a.
5 Thus the more physicalist naturalism of Hobbes or d’Holbach is very much a reduction to the physical properties of matter, while the naturalism of Gassendi or Diderot is a reduction to matter conceived as the bearer of vital, animate properties, typically attributed to minimal material components named ‘semences’, ‘seminalia rerum’, or ‘molecules’.
7 Walch 1726, cited and discussed in Rumore (ms. 2015).
Such mechanistic austerity (whether presented positively or negatively) stands in contrast to more dynamic materialisms like John Toland’s or Denis Diderot’s, where matter incorporates progressively more properties, from motion to thought, chemical dynamism, sensitivity and life. For Toland, “Activity ought to enter into the Definition of Matter, it ought likewise to express the Essence thereof”; “action is essential to Matter”; “I deny that Matter is or ever was an inactive dead Lump in absolute Repose, a lazy and unwieldy thing.”8 Diderot additionally attributes sensitivity9 to matter as a “universal” or “general” property: “All of matter senses … or tends to sense”; “if inert matter is arranged in a certain manner, impregnated with other matter, with heat and motion, it yields sensitivity, life, memory, consciousness, passions and thought.”10 Sometimes he presents this claim as deriving from chemistry, and describes the body as a chemical laboratory or distillation still, where sensitivity emerges: “The animal is the laboratory in which sensitivity shifts from being inert to being active.”11 But cerebral materialism exhibits an equally marked contrast between more active or dynamic and more passive or static understandings of the brain, which I shall discuss in the next section. Before proceeding further, however, I should like to point out one way in which the seeming sharpness of this distinction between materialism1 (the thesis of the materiality of the world) and materialism2 (a more ‘regional’ focus on mind-brain relations) can dissolve, which will help us grasp the novelty of Diderot’s version of materialism2 more clearly. Consider the case of physicalism.

If materialism was classically understood as materialism1 (everything that is real, is material or is the product of material processes), in the nineteenth and especially twentieth centuries, it primarily came to be understood as materialism2 (exploring the relation or ‘identity’ between mind and brain – between mental processes and cerebral processes). Now, both of these seem to indicate a privileged relation between materialism and scientific inquiry – or rather a privileged role for scientific inquiry. In the twentieth century, the science that predominated in this relation was physics. Materialism became synonymous with ‘physicalism’: the entities that were considered to be real – the basic facts of our ontology – were those described in, or ‘set by’ the physics of the time.

But what about the status of brains, including their ‘situation’ within a physicalist scheme? The identity theorists of the 1960s, notably J.J.C. Smart, were in fact quite satisfied with statements of physicalism, despite their proclaimed focus on mind-brain identity. Smart’s usage of Ockham’s razor in support of materialism is a physicalist parti pris: “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,”12 even if Smart does not propose some total reduction of neurophysiology to physics.13 Here is an uneasy alliance: is the materialist a brain theorist or a metaphysician bringing the rest of the world into line with physics? But how does one get from physics, or physicalism as an ontology, to the specific relation between brain states and mental

---

9 The French sensibilité is and was often translated ‘sensibility’ rather than ‘sensitivity’, but since in this context it refers to an organic property – the capacity to sense and respond to stimuli or impressions – rather than, say, moral sensibility, I use ‘sensitivity’. See overall Wolfe 2014b.
10 Diderot, Observations sur Hemsterhuis, in Diderot 1975-, XXIV, 278; Rêve, in Diderot 1975-, XVII, 105;
12 Smart 1959, 142, emphasis mine.
states? For Armstrong it was a “good bet.” In fact, it has been observed that much of the identity theory – in its blend of logic, semantics and ontological physicalism – conspicuously left out the messy details of neuroscience altogether – what Gary Hatfield called in a different context, the ‘neuroplumbing’. 

Contrasting both with such physicalist approaches to mind-brain identity and with a more general version of materialism (which he also develops elsewhere), the central figure in my discussion, Denis Diderot, is one of the first thinkers to notice, in writings of the 1760s-1770s, that any self-respecting materialist had to address the question of what brains do (including their ontology, i.e., their specific status in the causally specified natural world) and how much of our mental, affective, intellectual life is contained therein. This should not be confused with the prototypical cerebral-materialist move of attributing thought to the brain, which occurs at least as early as the 1700s with Toland, for whom “Whatever be the Principle of Thinking in Animals, yet it cannot be performed but by the means of the Brain” (he also explicitly describes thought as a property of the brain), and Anthony Collins, who defines consciousness as “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.” Neither Toland nor Collins feel compelled to provide more neuroanatomical detail or speculation about the details of our ‘neuroplumbing’; there is an implicit sense on their part that this is not part of the responsibility of the philosopher. Nor, unlike Diderot, do they consider that the ontological status of the brain is a key challenge in the articulation of a form of materialism, i.e. the brain as an organ possessing properties other than those of matter as a whole (or the body as a whole).

Doing justice to Diderot’s version of materialism also implies the recognition of a discontinuity, not just in forms of materialism overall, but between diverse materialist attitudes towards the brain – a diversity which is not just an effect of changing intellectual and socio-cultural contexts and practices, although I take seriously John Sutton’s suggestion that “it is possible to attend to contexts and to brains at once.” The specificity of Diderot’s view of the brain, I will argue, also reflects differing conceptions of the role of science, and the position of philosophical materialism with respect to such a role. If all defenders of materialism hold that mental and cerebral functions are correlated or identical, Diderot’s position is distinctive for three reasons.

First, his approach is less a prioristic than that of Toland, Collins or later, that of Bonnet and Priestley, in which ideas are movements of cerebral fibres (as discussed in the next section): he seeks to integrate an empirical dimension although without ruling out a speculative component (as discussed in the Conclusion). Conversely, his cerebral materialism is more sensitive to ontological implications than Thomas Willis’s neuroanatomy, which seeks to steer clear of such commitments. Second, Diderot’s portrayal of the brain-body-nervous system interface, via metaphors such as the harpsichord, is less mechanistic than what we think of as the Cartesian view (although contrast Sutton 1998 for a much more dynamic presentation of Descartes). Diderot also lacks the ‘localizationist’ impulse we might expect of a materialist, if it is true that all versions of materialism, conceptualize

14 Armstrong 1968, 90.
15 Bickle, Mandik, and Landreth 2010.
16 Hatfield 1992, 348.
17 Respectively, Toland 1704, letter IV, § 7, 139 and Toland 1720, 15.
19 Sutton 1998, 1. Projects such as ‘historical cognitive science’ or ‘historical neurophilosophy’ stress that theories of brain and cognition are historically diverse, as is cognition; see discussion in Present 2014.
the mind as “an ensemble of functions instantiated by spatially circumscribed and mutually connected portions of neural matter” (Métraux 2000, 183), with the implication that cerebral localization is “the point of intersection where mental functions are investigated naturalistically” (Young 1970/1990, vi).

The commitment to localize mental faculties in a portion of neural matter, which should not be confused with the broader claim that they are conceptually identical (as in Collins’s statement that consciousness is a quality inhering in the brain) itself predates the more specific claims of nineteenth-century neuroscience: earlier reflections on the “seat of the soul” were themselves attempts to ascertain a cerebral location of thought. Thus Robert Hooke wrote that “[T]here may be a certain Place or Point somewhere in the Brain of a Man, where the Soul may have its principal and chief Seat,” and added that this place is where “all the Impressions made from the Senses upon adapted Matter [are] delivered,” impressions which are “actual Locomotions given to the Parts of Matter or Bodies so or so moved”; the brain is a “repository.”20 As is well known, many authors, including Bonnet and Kant, wrestle with the ‘seat of the soul’, and often resort to invoking a kind of category mistake (thus Bonnet rather clumsily says the anatomy of the nerves shows that the soul is in the brain, but since the soul can have no ‘place’, “we will say the soul is present to the brain”21).

Thirdly, as an extension of the two first points, Diderot’s brand of materialism is distinctive because he acknowledges that the brain deserves special consideration, and recognizes a kind of plasticity, i.e., the fact that “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,”22 in a recent formulation. In his critique of Helvétius, Diderot clearly rejects a stimulus-response model of the relation between external sensory stimuli and the constitution cum individuation of our organism, and brain; he, too, is a species of determinist (Wolfe 2007) but one attentive to the malleability of organs like the brain. As I hope to show (section III), Diderot puts forth a form of materialism sensitive to cerebral plasticity, when he describes the brain as a book which reads and modifies itself.23 What does it mean for a materialist to recognize the ontological specificity of the brain, as notably conveyed in its plasticity, including in the most speculative sense, in which brains are ‘sculpted’ by their environments, “weak enough to yield an influence but strong enough not to yield all at once,” in William James’ definition of plasticity? That shall be my central question.

II. Cerebral materialism and plasticity: a prehistory

The question of materialism and plasticity can be seen to emerge from my second distinction (between passive and active versions of materialism) as applied to materialism, for brains can be presented as passive and static, or as dynamic and transformative, although the degree of complexity

20 Hooke 141, 140.
21 Essai analytique sur les facultés de l’âme, ch. V, § 27, in Bonnet 1771-1783, vol. VI, 13 (reprised in ch. XXIV, § 756 at 364); Bonnet insists this means he is not a materialist (ch. VIII, § 75, at 35). But, again clumsily, Bonnet also asserts that God needs only to vary brains, in his creations, to vary our souls (ch. XXV, § 771, at 370).
22 Malafouris 2010. On neuronal plasticity in historical perspective see Berlucchi and Buchtel 2009, and Huttenlocher 2002 for the contemporary discussions.
23 Éléments de physiologie, in Diderot 1975-, XVII, 470.
24 James 1914, 5-6; Berlucchi and Buchtel 2009, 307.
in a mechanistic portrayal of the brain can be quite advanced. More rudimentary early-eighteenth century versions such as Fontenelle’s present “the soul” as “think[ing] according to the material dispositions of the brain,” such that “certain motions in the brains” will yield “certain thoughts in the soul,” and conversely, “the objects of which we think leave material dispositions, i.e., traces in the brain.”

Hartley, Bonnet and Priestley will offer increasingly sophisticated versions of such views, correlating mental processes and e.g., cerebral vibrations, and sometimes seeking to work out the laws of correspondence between ideas and these vibrations and/or brain fibres.

Hartley ascribes “the performance of sensation” to “vibrations excited in the medullary substance,” quickly adding that this does not mean “that Matter can be endowed with the power of sensation”; Bonnet observes that even if we do not know how certain “brain fibres” produce “ideas in our soul,” nevertheless we know that without the movement of these fibres we would have no ideas; conversely, our brains are able to execute more or less complex ‘cognitive’ manoeuvres depending on the mobility of our brain fibres. Memory “belongs” to the brain for Bonnet, because our recalling of sensations depends on the communication of motions in the fibres. For Priestley, “as far as we can judge, the faculty of thinking, and a certain state of the brain, always accompany and correspond to one another”; thought belongs to a “particular organization” of matter, here the brain.

The obvious materialist implications do not worry Priestley (unlike Hartley or Bonnet): “the whole man is of some uniform composition”; “the property of perception, as well as the other powers that are termed mental, is the result (whether necessary or not) of such an organical structure as that of the brain”; more directly, thought “is a property of the nervous system, or rather of the brain.”

Yet whether they appeal to “traces,” “vibrations” or correspondences between ideas and cerebral fibres, these approaches, which treat mind and brain in primarily “procedural” fashion, do not take any form of plasticity into account. The most relevant feature of (neuro-)plasticity here is how it highlights the uniquely adaptive character of the brain. As John Sutton was the first to note (Sutton 1998, ch. 2), a very different vision of the brain can be found in the Oxford neuroanatomist Thomas Willis, whose writings are something of a conceptual hodge-podge, bringing together older theories of types of souls and newer concepts of animal spirits with a bold, unconstrained experimental-neuroscience program. On the traditional view, “Descartes’ physiology of the nervous system served as the foundation for all that has since been done in the interpretation of that system, and the modern view has in principle departed but little from the lead that Descartes gave it,” but recent historiography has given Willis pride of place. Indeed, if we distinguish between more ‘dynamic’ and more ‘static’ visions of the brain, Willis occupies a prominent place in the former category.

---

25 Fontenelle 1700/1818 (an anonymous work dated 1700, which first circulated in the 1743 Nouvelles libertés de penser), Part II.
26 Hartley 1749, I, 33.
27 Essai analytique, Preface, Bonnet 1771-1783, vol. VI, xii; ch. XVII.
28 Ibid., ch. XXV, § 793, at 380, and ch. XXI, 267-268; chapter XXII is primarily devoted to the brain.
29 Priestley 1777, 27, 28.
30 Priestley 1775, xx; Priestley 1777, 26, 27.
31 I cannot address ‘vibratory models of the nervous system’ in greater detail here, but a caveat to the above distinction would be that a degree of plasticity is arguably present once such models integrate metaphors such as the harpsichord (with nerves being like vibrating strings).
Willis describes complex cognitive processes such as memory and imagination in terms of the operation of animal spirits localised to different parts of the nervous system, largely based on comparative anatomy. His emphasis on dynamism and plasticity colours both his conception of the brain and of matter overall: the vision of matter as “meerly passive” is “vulgarly delivered” (Willis 1683, 33); rather, while acknowledging Gassendi who he admired, he develops a chimiatric concept of matter, composed of atoms which are “very active and self-moving” (ibid.). Matter here possesses various levels of chemical complexity and rarefaction, notably animal spirits, which are the result of complex processes of fermentation within the blood and the matter of the brain. Despite propelling the brain to the fore, however, Willis does not specifically worry about its specific ontological status, or its cultural embeddedness.

If we contrast Toland and Collins’ metaphysical assertions of mind-brain identity with Locke’s neat bracketing-off, at the beginning of the Essay, of cerebral analysis as irrelevant to an investigation of the mind (“I shall not at present meddle with the Physical consideration of the Mind”34), Willis conversely opens up neuroanatomy as an unbounded playing field, without committing to any dangerous metaphysical considerations on the nature of 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, including its relation to materialism, as discussed here. For the Willisian brain is clearly envisioned as a self-transforming, self-organising, plastic entity, while opponents – again, in an opposition we have encountered above – see the brain as more of a passive mechanism: according to John Hancock, in his 1706 Boyle Lecture attacking Willis35, the brain is 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 meer passive Principle, as to the Acts of inward Sensation and Intellection”; for Henry More, it was a mere “Cake of Sewet or Bowl of Curds,” a “little lurking Mushroom,” a “poor silly contemptible Knob” unfit to perform our cognitive operations.36 More insists rather predictably that “Brains have no Sense,” as they are not an active principle; additionally, the diversity of psychological faculties we experience (imagination, reason, memory, etc.) could not correspond to different parts of the brain, which would then be so many “Individual Persons” in the brain.37

Robert Boyle seems to be directly responding to More when he notes in the Christian Virtuoso 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...”38 Whether the issue is animal spirits, the irritability of muscles or sensitivity as a core property of the nervous system, the distinction between a model privileging transmission, adaptation and mobility, and one focusing strictly on brain architecture, correspondences or localization extends

---

33 CITE WILLIS ON SPIRITS
34 Locke 1975, I.i.2. For a different, provocative interpretation of empiricism as instead suffused with spirits, brain traces and other/materialities, see Sutton 2010. That Locke was also Willis’s student does not seem particularly relevant to Locke’s discussion of ideas and their relation (or lack thereof) to the brain, although he owes, at minimum, some of his critique of innate ideas to his teacher.
35 Hancock 1739, II, 243.
37 More 1653, 37, 39; 38.
38 Boyle 1772/1966, VI, 741.
well beyond early modern England! Boyle’s assertion contra More, that there must be more to the brain “than is obvious to a vulgar eye” (even “that of a dissector”), can indeed be generalized all the way to nineteenth-century localizationist neuroscience, which neglects or factors out the malleability and adaptive properties of the brain.

Of course, to point to the existence of “far more” cerebral complexity than is recognized by anti-materialists who target materialism can itself be quite consistent with a (complex) mechanistic picture of the brain, e.g. the various programs which seek to map out cognitive activity by correlating (or identifying) components of cognition with components of cerebral architecture. This focus on correlation, quantification and localization – as in, e.g., Bonnet’s project to detail how different impressions affecting the mind differ in vivacity according to the “intensity of motions communicated to the brain fibres” and “the vivacity of sensations is necessarily proportional to the intensity of the motions that excite them” – should not be confused with theories of mind and brain which focus on ‘imprinting’, where the emphasis is on how mental activity (including influences from culture, mediated by the imagination) can imprint material traces on the brain; but a model of cerebral traces is not in itself a model of plasticity.

Curiously, this acknowledgement of ‘imprinting’ is not initially due to experimentalists. Somewhat unexpectedly, it is Locke, and then Laurence Sterne who recognize the cerebral materiality involved, not just in the regular processes of mental life, but in its plasticity. This is manifest in Locke’s often-unnoticed comment that forms of mental pathology which are usually traced back to behaviors (e.g. being frightened in childhood by stories of goblins), are in fact co-constituted by such behaviors and by their imprint in the brain, qua “Trains of Motion in the Animal Spirits” One can contrast Locke’s account here with a more Cartesian conception of self-perfection (precisely a form of malleability) wherein, as Pieter Present puts it, “a human endowed with reason can train the dog’s brain, but a brain cannot make itself more reasonable”; “a brain cannot make itself more perfect” (Present 2014, § 3.3.2), i.e. there is Cartesian plasticity, but not of a materialist sort. The brain, as material, is necessarily passive, and the immaterial self organizes the brain, whereas for Diderot, as we will see below, brain plasticity and the activity of matter enable a vision of the self-organizing brain (ibid.). Granted, Locke is not the most obvious candidate for early intimations of brain plasticity, given his explicit “bracketing-off” of naturalistic considerations concerning the brain and mental activity which position the Essay as a non-materialist work: as noted, he will not provide “physical consideration of the mind” (Locke 1975, I.i.2), including “what motions of our spirits or alterations of our bodies we come to have any sensation by our organs, or any ideas in our understandings; and whether those ideas do in their formation, any or all of them, depend on matter or not” (ibid.).

But Locke is not the only non-materialist author to hint at the receptivity and malleability of the brain. Another such intimation of plasticity, closer to Diderot, is Sterne’s Tristram Shandy, a work

---

39 Essai de psychologie, ch. VII, in Bonnet 1771-1783, vol. VIII, 13; Essai analytique, in Bonnet 1771-1783, vol. VI, ch. XI. 79. Timo Kaitaro suggests, in an elegant paper which complements this one, that the identity theory really matches Bonnet’s dualism rather than Diderot’s materialism. He points out 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(s) in the brain for each idea. However, “the metaphysical interpretation of these identities depends on whether one is a materialist or a dualist, but on the basis of the historical analysis of localizationist doctrines … that the postulation of such identities in itself is not committed to dualism or materialism” (Kaitaro 2004, 629).

40 Locke 1975, II.iii.6 (CITE LONGER PASSAGE); see Sutton 2010 for further discussion.
which influenced Diderot’s prose endeavours. *Tristram Shandy* is famous, among other reasons, for its literary appropriation and usage of animal spirits, which yields a dynamic materialism of matter-mind relations. But the title character also describes an associative mechanism whereby, when arriving in Lyon, he cannot help thinking about a story of two Lyonnais lovers he had read in his youth, “a sweet æra in the life of man, when (the brain being tender and fibrillous, and more like pap than anything else…)”; he cannot avoid this recollection because the cultural-associative event of reading about them had modified his brain anatomy: “There is a soft æra in every gentle mortal’s life, where such a story affords more *pabulum* to the brain, than all the *Frusts*, and *Crusts*, and *Rusts* of antiquity, which travellers can cook up for it.”

If we recall my broad distinction between static (passive) and dynamic (active) conceptions of the brain, the diverse models of the brain ‘on offer’ or in the process of articulation in this period display a specific version of the distinction, namely, between more mechanistic, proto-localizationist analyses of interconnections (which search for correlations and/or identifications of, e.g., ideas and cerebral fibres) and more ‘imprinting’-oriented models (which seek to relate the dynamism of mental life, as in imagination or hallucinations, and changes in brain architecture), as in Locke and Sterne. Searching for structures and correlations is different from intimations of plasticity, including cultural plasticity, as I discuss below. Diderot will extend such intimations further, emphasizing in addition that the brain is self-organizing, and that its description is interdependent with a new version of materialism.

**III. “The brain is a book which reads itself”: Diderot’s plastic-cerebral materialism**

Diderot’s reflection on brains, minds, nerves and plasticity occurs across a variety of writings, including novels, scientific commentary, and plain ‘philosophy’, as in the *Rêve de D’Alembert* (1769; unpublished), which introduces scientific speculations with metaphysical ramifications. But it is in his late manuscript on ‘physiology’, in fact a kind of natural-philosophical handbook for materialism, the *Éléments de physiologie* (written in the 1770s-1780s and unpublished42), that we find the central passage for my analysis. It occurs in a chapter on memory, in the third and last section of the manuscript, dealing with “phenomena of the brain.” Diderot presents several extremely lyrical cases of recalling landscapes in nature and landscapes in painting, and then almost abruptly turns to cerebral-material explanations of such phenomena, with a striking image:

> In order to explain the mechanism of memory we have to treat the soft substance of the brain like 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.43

---


42 It is not clear if Diderot intended to publish the *Eléments* or not, or thought of it as a finished work. Its title is taken from Albrecht von Haller’s influential textbook *Elementa Physiologiae* (6 vols., 1757-1766), but it reflects a variety of influences and medico-clinical sources. As Warman 2012 notes, none of the three scholarly editions that exist, Mayer’s 1964 edition, his newer 1987 edition (in Diderot 1975-., vol. XVII), or Quintili’s 2004 edition have any commentary on this passage.

43 *Éléments de physiologie*, in Diderot 1975-., XVII, 470.
There are prefigurations of this image of the brain as a book. In his 1755 *Essai de psychologie*, Charles Bonnet speaks of the “mechanism of the brain” which could be read by an “Intelligence” which would be familiar with all of its details, sounding rather like Laplace’s demon some decades later; Bonnet adds that that this intelligence would read the brain “like a book.” This is just one step removed from Diderot’s conception (in a sense, just like Locke and Sterne, who recognized imprinting but not self-organization). But a less-known author may also have influenced Diderot here: the heterodox Benedictine monk Léger-Marie Deschamps, who met with Diderot several times during the summer of 1769, when Diderot was writing the *Rêve*; Deschamps showed Diderot his ‘clandestine’ materialist treatise, *La Vérité ou le vrai Système*, which sets out a Spinozist metaphysics of relations, including a discussion of the interrelation of sensation and the world of objects. Deschamps describes how “to read me, to hear me read is to become composed of (se composer de) my work, which then acts physically by the eyes or ears on the brain fibres, and raises them to a given tone, according to its impressions on them.” Earlier in the century, Fénelon described the brain as “a kind of book” filled with an almost infinite number of images and characters (CITE). Here the context is not at all materialist, but Diderot’s version employs all of these nuances to articulate a unique form of materialism.

I have already contrasted Diderot with Bonnet and others, but consider what an author much closer to Diderot, namely La Mettrie, has to say about the brain: where Diderot emphasized the receptivity of the “soft substance of the brain,” La Mettrie does not. He either remains at a level of greater generality (“all the faculties of the soul depend to such a degree on the brain and the whole body’s own organization that they visibly are nothing but this organization itself”) or he repeats older ideas of passive imprinting of traces. Much the same contrast is apparent in their respective uses of the image of the harpsichord to explain the nervous system. La Mettrie had compared the brain to a harpsichord, with sensitive vibrating chords that form a totality unified by imagination (a system of interlocking and reverberating fibres, i.e. chords: sounds enter through the ear, images through the eye, and strike various chords; add memory and one has the three cognitive faculties). But he does not use this analogy in support of a broader analogy-based conception of both the nervous system and human organic life overall, as in Diderot’s “We are instruments endowed with sensitivity…” Bonnet uses the image too, but in a more dismissive way, to convey what happens when we associate ideas in a mad, haphazard way: our brain is then like a harpsichord whose keys are touched by “an ignorant hand.” When Diderot uses the analogy, it resonates much more, so to speak, with the image of the brain as book which reads itself. In an attempt to describe the interconnected, ‘systemic’ nature of the nervous system and the property of sensitivity, he writes that

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.

44 Bonnet 1771-1783, vol. VIII, 2.
45 Dom Deschamps remains a mysterious figure; his influence on Diderot was first noted by Vernière in his edition of Diderot’s works (Diderot 1961, 300, n. 1).
46 Deschamps 1993, 404; 385.
47 Fénelon 1713, § XLIX, 168-170.
We are instruments endowed with sensitivity and memory; our senses are so many keys that are struck by surrounding nature, and that often strike themselves…

Diderot sees that a concept such as sensitivity allows him to conceptually integrate the reactivity and 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. Yet these ‘network’ properties are not properties of matter as such, but of a body-brain network. Specifically, sensitivity is the property of the network, as described in a series of metaphors, including the harpsichord, the spider and its web, and a crawfish. That the harpsichord-model of sensitivity implies that we ourselves are like “sensing instruments,” is, again, not unlike the image of the brain as a self-reading book, and this in fact implies that Diderot is integrating what might seem like two different perspectives: an emphasis on the primacy of the nervous system (in the Rêve) and a more “cerebrist” emphasis on the brain in the Éléments. Diderot in fact analyses both as belonging to a body-brain network (like the spider and the spiderweb, which he views as forming one organic system), which he describes as “a system of actions and reactions.”

This interest in sensitivity should also be understood as underscoring how we are not like the passive recording mechanisms which fascinated earlier generations of natural philosophers: as Diderot exclaims, “What a difference there is, between a sensing, living watch and a golden, iron, silver or copper watch!” Notably, “our key characteristics lie in our brains, not in our external constitution (organisation, a term used in this context to mean our overall physiological configuration)” so that “in order to explain the mechanism of memory we have to examine the soft substance of the brain.” A cognitive property such as memory is a product of our organisation: “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. This is different from Hobbes or Collins’ materialism, to be sure, 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. Wasn’t this also La Mettrie’s view, and that of Bonnet and Priestley (with admittedly different neuroarchitectures in each case)?

What is different is Diderot’s central choice of metaphor for the brain, to quote it again. First, “the soft substance of the brain” is “like 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.” Second, he calls this “the book” (itself not a new image, for descriptions of memory as being like an imprint on a tablet go back at least as far as Plato’s Timaeus: Warman 2012), but a book which has as

---

51 Diderot, Rêve de D’Alembert, in Diderot 1975-, XVII, 102.
52 Diderot, Éléments de physiologie, in Diderot 1975-, XVII, 355.
53 I thank Ann Thomson for making me clarify this point (and indeed the crawfish image is both a network-of-sensibility image, and one which leads back to the brain). On brains and nervous systems in this period, see Laplassotte 1970, 609.
54 Diderot, Éléments, in Diderot 1975-, XVII, 337.
55 Diderot, Éléments, 335.
56 Diderot, Éléments, 326, 470.
57 Rêve de D’Alembert, in Diderot 1975-, XVII, 101; Éléments, 335. Elsewhere Diderot also attributes ‘intentional’ properties to the tiniest components of living matter, in a kind of pan-psychism (although he criticizes Maupertuis for just this in his Pensées sur l’interprétation de la nature: see Wolfe 2010).
its reader, “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.” But the wax tablet is not a self-reading tablet, just as the brain as a Renaissance memory chamber is not reading itself! And if one looks forwards rather than backwards to Plato or the Renaissance, one can also say that Diderot is not seeking to materialize memory in an ‘engram’ concept, i.e., a material location of memory storage (Logan 2015). Rather, he uses the case of memory as a pretext for highlighting the ‘Jamesian’ plasticity of the brain: “weak enough to yield an influence but strong enough not to yield all at once.”

In fact, Diderot seems to move between different positions concerning the brain, even in the Éléments de physiologie: either our key characteristics are located in our brain (thus the brain is special, perhaps ontologically unique), or the brain is just “an organ like any other,” “a secondary organ” (467), or “merely a secretory organ” (353). If the brain is ‘an organ like any other’, it is notably because its complex properties would themselves be dependent on the nervous system and more metaphysically, on sensitivity as a property of matter. Obviously, the brain as a book which reads itself is an instance of the former view, like Diderot’s comment elsewhere, that the brain is like the “judge” to which the five senses report as “witnesses”: “there is a particular organ, the brain, to which the five witnesses report. This organ deserves particular study”; this is why “Man’s key characteristics lie in his brain, not in his external constitution.”

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 a degree of ‘anthropocentrism’ within a naturalistic framework: that is, if the human 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. Indeed, whether the brain is special or not, it belongs to a network model of sensitivity, which itself determines a metaphysics of living, sensing matter. The “mass of sensitive and living wax,” i.e. self-organizing matter possessing the property of sensitivity, is closely linked to Diderot’s conception of a universally sensing matter.

The brain as a book which reads itself is different from other material arrangements and by extension from traditional mechanisms, much as “living, sensing watches” differ from “copper or iron watches.” Diderot seems to collapse any distinction between materialism1 (an ontology of matter and its properties, such as motion, irritability, sensitivity, thought) and materialism2 (‘cerebral materialism’ focusing on brain-mind identity) , while emphasizing activity and dynamism. The difference between these two general types of materialism is collapsed, implying that they can also be combined: a dynamic matter theory can ‘facilitate’ a dynamic picture of cerebral plasticity (in Diderot, they seem to relate in that order). After all, this is what some opponents of materialism warned against early on, as in Cudworth’s (worried) assertion that “It is demonstrably evident and mathematically certain, that no Cogitation can possibly arise out of the Power of Matter.”

Conclusion

58 Éléments, 470.
59 James 1914, 5-6.
60 Respectively, Réfutation d’Helvétius, II, ch. xii, in Diderot 1975-, XXIV, 549 and Éléments, 326.
61 Cudworth 1731, 302.
Different forms of materialism diverge as they address the status of the brain, some developing more *plastic*, culturally embedded models of the brain, some more formal, mechanistic models. Here, the difference is not between specifically cerebral materialism versus larger claims about fundamental physics or logical identities: Toland, Collins, Priestley and even Bonnet also explicitly identified cerebral processes and mental processes, like nineteenth-century localizationists (although such identifications are not necessarily *philosophically* materialist). Diderot’s form of materialism is not focused on such identities, including the project of localizing mental functions in specific brain areas. It is rather a *cultured-brain* materialism, in which the brain is not a passive recording machine like More’s “poor contemptible Knob” or “rude lump of soft matter,” but rather, a plastic mass of “sensitive and living wax,” which is self-organizing, self-interpreting and co-constituted in relation to the external world. Indeed, plasticity does not imply that the brain develops on its own by a kind of Baron von Münchhausen manoeuvre: as Diderot observes, “objects are required for the brain to think” (Diderot 1975-, XVII, 466). But defining the brain as a book which reads itself has one further implication I would like to emphasize, regarding the relation between materialism and science, or in this case scientism.

If one thinks back to mind-brain identity claims as ‘empirical’, it seems quite natural to imagine that the materialist philosopher would build on these as evidence, going so far as to ultimately factor out philosophy. But Diderot’s conception of the brain is not so fatal to philosophy, unlike some identity-theoretic proclamations such as U.T. Place’s:

[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 has happened in the lifetime of most of us in the case of the origins of the universe which 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.\(^62\)

Here, the theory of brain-mind identity moves closer to the status of a verifiable or falsifiable scientific theory, as “brain-imaging evidence begins to replace the subject’s introspective report in determining the occurrence and nature of her conscious experience.”\(^63\) It is a much more sophisticated identity theory, but philosophically it is akin to Carl Vogt’s infamous identity-slogan of the nineteenth century (in fact reprising Cabanis), “thought is to the brain what bile is to the liver or urine to the kidneys.”\(^64\)

Diderot’s materialism is not of this sort, nor does materialism in general *have to be* a ‘handmaiden’ of the natural sciences, both because its claims are not necessarily founded on experimental evidence (the brain as the book which reads itself is a metaphor) and because it sometimes takes a deliberately speculative form. In this sense, even if Diderot was seeking to understand the ‘neuropsychology’ of memory, he nevertheless illustrates a more general point that materialism need not be a scientism. If purely conceptual, a priori arguments for mind-brain identity can be made (as in Toland and Collins), claims like Place’s, that mind-brain identities are objects of experimental study, soon to pass from the philosophical “grasp,” lie at the opposite end of the

---

\(^62\) Place 1997, 16. Thanks to Dr. M.-C. Wright (University of Leeds) for providing me with a copy of Place’s paper.

\(^63\) Place 1997, 15.

\(^64\) Vogt 1847/1874, XIII, 323.
The danger of scientism lies in confusing conceptual (here, *philosophically* materialist) and empirical (scientific) claims. Vogt’s ‘vulgar’ statement of brain-mind identity is not problematic as such: it seeks to correlate and/or identify mental processes and cerebral processes. The problem lies in confusing genres, when materialism is presented as itself a straightforward empirical claim, unlike Diderot’s version of materialism, which is not solely founded on experimental evidence.

But then we encounter again the problem I raised above: is there something anthropocentric about Diderot’s cerebral-materialism? If the human brain is a “special organ” rather than an organ like any other, this may reflect a desire on the part of a materialist philosopher to retain a modicum of autonomy for humans (actually, higher mammals overall); what Diderot elsewhere calls the “presence of man” within a naturalistic universe, a presence without which “this moving and sublime spectacle of nature becomes nothing but a sad and mute scene. ... a vast solitude.” Yet rather than inflate these indications into a form of humanism that would be either incoherent with materialism or a kind of exception, I think more minimally that the question we should retain is, can a materialist hold that the brain is special? and the answer is 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. It is not just the property of sensitivity which differentiates certain material arrangements from others, but the additional feature of self-organization, reflecting an inherent plasticity, open to the cultural and social determinations of brain activity, and thereby, as Diderot says, the real basis of our individuality (not our external conformation).

While for other types of materialism, such as the identity theory, there is physics, and anything above (both biology and neuroscience) is like special kinds of radio-engineering, in contrast, for Diderot the brain is ontologically specific: he was concerned with real, flesh-and-blood brains and their specificities. By acknowledging that the brain (and nervous system) have particular properties distinct from physical nature as a whole, he grants it a special ontological status. But can materialism maintain that the brain has an ontology without reintroducing special laws and properties of human nature, different from the laws of nature as a whole (‘kingdoms within kingdoms’, in Spinoza’s formulation)? The question is still an open one today.

Acknowledgments
Thanks to Pieter Present for his comments on an earlier version of this paper

References


---


Bonnet, Charles. 1771-1783. Œuvres d’histoire naturelle et de philosophie, 8 tomes (some in 2 volumes). Neuchâtel: S. Faulche.


More, Henry. 1653. An antidote against atheisme, or, An appeal to the natural faculties of the minde of man, whether there be not a God. London: Roger Daniel.


Place, U.T. 1997 ms. “We needed the Analytic-Synthetic Distinction to formulate Mind-Brain Identity then: we still do.” Symposium ‘40 Years of Australian Materialism’, Dept. of Philosophy, University of Leeds.


Rumore, Paola. 2015 (ms.). “Mechanism and Materialism in Early Modern German Philosophy.”


Toland, John. 1720. Pantheisticon, sive formula celebrante sodalitatis socratiae… s.n.: ‘Cosmopoli’.


Warman, Caroline. ms. 2012. Variations on the figure of the mind as book: Diderot’s mécanisme de la mémoire.


